Accurate Calculations of Molecular Properties with Explicitly Correlated Methods

Jinmei Zhang

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Chemistry

Edward F. Valeev, Chair
T. Daniel Crawford
George A. Hagedorn
Louis A. Madsen
Diego Troya

July 28, 2014
Blacksburg, Virginia

Keywords: Quantum Chemistry, Explicitly Correlated Methods, High Accurate Computation, Reaction Barriers, Thermochemical Properties

Accurate Calculations of Molecular Properties with Explicitly Correlated Methods

Jinmei Zhang

(ABSTRACT)

Conventional correlation methods suffer from the slow convergence of electron correlation energies with respect to the size of orbital expansions. This problem is due to the fact that orbital products alone cannot describe the behavior of the exact wave function at short inter-electronic distances. Explicitly correlated methods overcome this *basis set* problem by including the inter-electronic distances (r_{ij}) explicitly in wave function expansions. Here, the origin of the basis set problem of conventional wave function methods is reviewed, and a short history of explicitly correlated methods is presented. The F12 methods are the focus herein, as they are the most practical explicitly correlated methods to date. Moreover, some of the key developments in modern F12 technology, which have significantly improved the efficiency and accuracy of these methods, are also reviewed.

In this work, the extension of the perturbative coupled-cluster F12 method, CCSD(T)$_{\overline{\text{F12}}}$, developed in our group for the treatment of high-spin open-shell molecules (J. Zhang and E. F. Valeev, *J. Chem. Theory Comput.*, 2012, **8**, 3175.), is also documented. Its performance is assessed for accurate prediction of chemical reactivity. The reference data include reaction barrier heights, electronic reaction energies, atomization energies, and enthalpies of formation from the following sources: (1) the DBH24/08 database of 22 reaction barriers (Truhlar et al., *J. Chem. Theory Comput.*, 2007, **3**, 569.), (2) the HJO12 set of isogyric reaction energies (Helgaker et al., *Modern Electronic Structure Theory*, Wiley, Chichester, first ed., 2000.), and (3) the HEAT set of atomization energies and heats of formation (Stanton et al., *J. Chem. Phys.*, 2004, **121**, 11599.). Two types of analyses were performed, which target the two distinct uses of explicitly correlated CCSD(T) models: as a replacement for the basis-set-extrapolated CCSD(T) in highly accurate composite methods like HEAT and as a distinct model chemistry for standalone applications. Hence, (1) the *basis set error* of each component of the CCSD(T)$_{\overline{\text{F12}}}$ contribution to the chemical energy difference in question and (2) the *total error* of the CCSD(T)$_{\overline{\text{F12}}}$ model chemistry relative to the benchmark values are analyzed in detail. Two basis set families were utilized in the calculations: the standard aug-cc-p(C)VXZ (X = D, T, Q) basis sets for the conventional correlation methods and the cc-p(C)VXZ-F12 (X = D, T, Q) basis sets of Peterson and co-workers that are specifically designed for explicitly correlated methods. The conclusion is that the performance of the two families for CCSD correlation contributions (which are the only components affected by the explicitly correlated terms in our formulation) are nearly identical with triple- and quadruple-ζ quality basis sets, with some differences at the double-ζ level. Chemical accuracy (~ 4.18 kJ/mol) for reaction barrier heights, electronic reaction energies, atomization energies, and

enthalpies of formation is attained, on average, with the aug-cc-pVDZ, aug-cc-pVTZ, cc-pCVTZ-F12/aug-cc-pCVTZ, and cc-pCVDZ-F12 basis sets, respectively, at the CCSD(T)$_{\overline{\text{F12}}}$ level of theory. The corresponding mean unsigned errors are 1.72 kJ/ mol, 1.5 kJ/mol, $\sim$ 2 kJ/mol, and 2.17 kJ/mol, and the corresponding maximum unsigned errors are 4.44 kJ/mol, 3.6 kJ/mol, $\sim$ 5 kJ/mol, and 5.75 kJ/mol.

In addition to accurate energy calculations, our studies were extended to the computation of molecular properties with the MP2-F12 method, and its performance was assessed for prediction of the electric dipole and quadrupole moments of the BH, CO, H_2O, and HF molecules (J. Zhang and E. F. Valeev, in preparation for submission). First, various MP2-F12 contributions to the electric dipole and quadrupole moments were analyzed. It was found that the *unrelaxed* one-electron density contribution is much larger than the orbital response contribution in the CABS singles correction, while both contributions are important in the MP2 correlation contribution. In contrast, the majority of the F12 correction originates from orbital response effects. In the calculations, the two basis set families, the aug-cc-pVXZ (X = D, T, Q) and cc-pVXZ-F12 (X = D, T, Q) basis sets, were also employed. The two basis set series show noticeably different performances at the double-ζ level, though the difference is smaller at triple- and quadruple-ζ levels. In general, the F12 calculations with the aug-cc-pVXZ series give better results than those with the cc-pVXZ-F12 family. In addition, the contribution of the coupling from the MP2 and F12 corrections was investigated. Although the computational cost of the F12 calculations can be significantly reduced by neglecting the coupling terms, this does increase the errors in most cases. With the MP2-F12$_{\text{C}}$/aug-cc-pVDZ calculations, dipole moments close to the basis set limits can be obtained; the errors are around 0.001 a.u. For quadrupole moments, the MP2-F12$_{\text{C}}$/aug-cc-pVTZ calculations can accurately approximate the MP2 basis set limits (within 0.001 a.u.).

This work received the financial support of the U.S. National Science Foundation.

Acknowledgments

First, I would like to thank Dr. Edward Valeev for his guidance in my research and his dedication to help me improve both my research and communication skills. I also like to thank my committee members, Dr. Daniel Crawford, Dr. George Hagedorn, Dr. Louis Madsen, and Dr. Diego Troya, for their support and guidance throughout my graduate school years. Finally, I would like to thank my husband, Justus Calvin, for his support during the years and hard work for reading and editing my thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Through its eighty-year history, quantum chemistry has developed into a useful tool for chemists in a wide variety of areas to explain experimental observations, validate experimental data, and even predict experimental results. In some applications, computation has become as reliable as experiments. Wave function methods are one of the important components in quantum chemistry. Despite the success of these methods, the conventional wave function methods suffer from the slow convergence problem of electron correlation energies in orbital expansions, and this problem is exacerbated by the steep increase of computational cost with the basis set size (i.e. as the orbital expansion becomes larger).

The slow convergence problem is due to the sole use of Slater determinants in conventional wave function methods, which are written as products of orbitals (one-electron functions). However, orbital products alone cannot describe the behavior of the exact wave function at short inter-electronic distances, where it has a cusp. In standard wave function methods, such as the coupled-cluster singles and doubles augmented by perturbative treatment of triples (CCSD(T)) or multireference configuration interaction (MRCI) methods, basis sets errors of electron correlation energies in atoms decrease as $\mathcal{O}[(L_{\max}+1)^{-3}]$ for a basis set saturated to the angular momentum $L_{\max}$.[8] This problem can be overcome by using explicitly correlated methods, which include the inter-electronic distances (r_{ij}) explicitly in the wave function expansions. For atoms, the explicitly correlated methods proposed by Kutzelnigg,[9] commonly known as R12, or F12, methods, have a basis set error of $\mathcal{O}[(L_{\max}+1)^{-7}]$.

The efficiency and accuracy of F12 methods have been significantly improved in the past decade by the modern F12 technology, and they have become common tools in computational chemistry to calculate accurate ground-state energies.[2,10] In general, modern F12 methods require a basis set of two cardinal numbers lower than the comparable conventional methods in order to obtain the same accuracy. In the past, the F12 methods have been mainly used to obtain a better description of correlation energies, whereas F12 calculations of molecular properties, such as dipole moments and polarizabilities, can be found in only a few studies.[11–14] Compared with the single-point energy calculation, the analytic evalu-

ation of molecular properties is much more complicated and computationally demanding. The extension of F12 methods to calculations of molecular properties offers the possibility of their accurate prediction for larger systems with lower computational cost. Furthermore, F12 calculations cannot be considered as black-box applications yet, despite the tremendous progress in the development of F12 methods. One of the issues, among many, is the requirement for particular orbital basis set types in F12 computations. So far, the recommended basis sets remain the cc-pVXZ-F12 series of Peterson et al.[15–17] and the standard aug-cc-pVXZ series of Dunning et al.[18–20] One of our goals in this work is to systematically evaluate these two basis set series for F12 calculations of various molecular properties.

In this work, we first present a brief review of conventional wave function methods, and then discuss the origin of the slow convergence problem in these methods. A short history of early explicitly correlated methods is given in Chapter 2, but our focus is the modern F12 methods (Chapter 3), the most practical and promising explicitly correlated methods. In Chapter 4, we document the extension of the perturbative coupled-cluster F12 method, CCSD(T)$_{\overline{\text{F12}}}$, developed in our group for the treatment of high-spin open-shell molecules. We assessed its performance for accurate studies of chemical reactivity by performing benchmark calculations of reaction barrier heights and thermochemical properties including electronic reaction energies, atomization energies, and heats of formation. In the last chapter, we present the analytical evaluation of the relaxed one-electron density with the MP2-F12 method, and discuss its application to compute the one-electron molecular properties: electric dipole and quadrupole moments.

1.1 Born-Oppenheimer Approximation

The central equation in quantum chemistry is the time-independent molecular Schrödinger equation,

$$\left(-\frac{1}{2}\sum_i \nabla_i^2 - \frac{1}{2M_A}\sum_A \nabla_A^2 - \sum_i\sum_A \frac{Z_A}{r_{iA}} + \sum_{A<B}\frac{Z_A Z_B}{r_{AB}} + \sum_{i<j}\frac{1}{r_{ij}}\right)\Psi = E\Psi \qquad (1.1)$$

where i and j represent the ith and jth electrons, A and B represent the Ath and Bth nuclei, Z is the nuclear charge, and r is the distance between two particles. This equation describes the physics of the electrons and nuclei in the atoms and molecules. If we could find the solution to this equation, we would be able to determine all properties of the system. Unfortunately, we can only solve this equation exactly for one-electron systems (hydrogen-like atoms) as the potential energy of the electron-electron interaction ($\sum r_{ij}^{-1}$) couples the coordinates of all electrons. For all other systems, we can only obtain approximate solutions to this equation. In most chemistry applications, the Born-Oppenheimer (BO) approximation[21] is assumed, where only the motion of electrons is considered since nuclei are much heavier than electrons, and thus move much slower. The BO approximation simplifies the

problem to solving the electronic Schrödinger equation,

$$\left(-\frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_A \frac{Z_A}{r_{iA}} + \sum_{i<j} \frac{1}{r_{ij}}\right) \Psi_{\text{elec}} = E_{\text{elec}} \Psi_{\text{elec}}, \tag{1.2}$$

where the kinetic energies for the nuclei are zero, and the nuclear repulsion becomes constant. Even within this framework, however, it is still impossible to find the exact solution to the electronic Schrödinger equation of many-electron systems.

1.2 Hartree-Fock Method

The Hartree-Fock (HF) method[22] is one of the simplest approximate methods used to solve the electronic Schrödinger equation. Moreover, it often serves as the starting point for other, more accurate methods. For the ground state of an N-electron system, the HF wave function is a single Slater determinant of molecular spin-orbitals,

$$\Psi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}, \tag{1.3}$$

where a molecular spin-orbital occupied by an electron, $\chi(\mathbf{x})$, is comprised of two parts: a spatial function, $\psi(\mathbf{r})$, and a spin function, $\alpha(w)$ or $\beta(w)$, and $\mathbf{x}$ represents both the spacial and spin coordinates: $\mathbf{x} = \{\mathbf{r}, w\}$. For convenience, the Slater determinant is often written as

$$\Psi_0 = |\chi_1 \chi_2 \cdots \chi_N\rangle, \tag{1.4}$$

and the spin-orbital, $\chi(\mathbf{x}_i)$, is usually written as $\chi(i)$. In most molecular calculations, the spatial function is expressed as a linear combination of basis funcitons,

$$\psi_i(\mathbf{r}) = \sum_{\mu=1}^{K} C_{\mu i} \phi_\mu(\mathbf{r}) \quad (i = 1, 2, \cdots, K) \tag{1.5}$$

where $\{\phi_\mu\}$ is a set of known functions (atomic orbitals or basis functions), and $C_{\mu i}$ are molecular orbital (MO) coefficients which can be determined using the variational method. This approach is known as the Hartree-Fock-Roothaan method.[23] In this method, the precision of the HF method is determined by the completeness of the basis set, i.e. the more complete the basis set, the more precise the result. Within the HF approximation, a complete basis set yields the exact solution, the Hartree-Fock limit. In practice, however, finite basis sets are used, which introduces errors in the energy.

The HF energy is obtained by minimizing the expectation value of the Hamiltonian, $\langle\Psi_0|\hat{H}|\Psi_0\rangle$, with respect to the molecular orbitals. In this process, the following Hartree-Fock equations can be obtained:

$$\hat{F}_i\chi(i) = \varepsilon_i\chi(i) \tag{1.6}$$

with

$$\hat{F}_i = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + v_{\mathrm{HF}}(i), \tag{1.7}$$

where $\hat{F}_i$ is the Fock operator, and $v_{\mathrm{HF}}(i)$ is the HF potential, an average potential experienced by the ith electron. The HF potential can be written as

$$v_{\mathrm{HF}}(i) = \sum_j \left(\hat{J}_j(i) - \hat{K}_j(i)\right), \tag{1.8}$$

where the Coulomb operator, $\hat{J}_j$, and the exchange operator, $\hat{K}_j$, are defined as

$$\hat{J}_j(1)\chi_i(1) = \langle\chi_j(2)|\frac{1}{r_{12}}|\chi_j(2)\rangle\chi_i(1), \tag{1.9}$$

$$\hat{K}_j(1)\chi_i(1) = \langle\chi_j(2)|\frac{1}{r_{12}}|\chi_i(2)\rangle\chi_j(1). \tag{1.10}$$

Thus, the HF potential is comprised of two parts: the average Coulomb potential and the exchange potential which results from the antisymmetric requirement of the wave function with respect to the interchange of electronic coordinates. Essentially, the spontaneous electron repulsion, or correlation, is replaced by an average potential field in the HF method. The difference between the exact (nonrelativistic) energy and HF energy of a system is called the electron correlation energy:

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}. \tag{1.11}$$

While the HF energy typically accounts for more than 99% of the total electronic energy of a molecule, the electron correlation energy, is critical for the accurate prediction of quantities of chemical interest. Consider the binding energy in H_2 as an example: the contribution from the correlation effects is about 25 kcal/mol.[24] As the number of electron pairs increases, so do the correlation effects. For highly accurate calculations, it is necessary to properly describe the correlation between the motion of the electrons.

1.3 Standard Correlation Methods

In a HF calculation with a basis set of K functions, we can obtain $2K$ spin-orbitals: N occupied orbitals and $2K-N$ virtual (or unoccupied) orbitals, where N is the number of electrons in the system. While the HF ground state, Φ_0, is formed by occupied spin-orbitals, excited

determinants, $\Phi_{ij\ldots}^{ab\cdots}$, can be formed by substituting occupied orbitals (denoted by i, j, ...) in Φ_0 with virtual orbitals (denoted by a, b, ...). In the conventional correlation methods,[25] such as configuration interaction (CI) methods, coupled-cluster (CC) methods, and many-body perturbation theory (MBPT), electron correlation effects are taken into account by including these excited determinants in the approximate wave functions. The differences among these methods lie in the computational procedures for determining expansion coefficients and the expressions for molecular energies and properties.

1.3.1 Configuration Interaction methods

In CI methods,[22] the approximate wave function is written as a linear combination of multiple determinants:

$$\Psi = C_0\Phi_0 + \sum_{i,a} C_i^a \Phi_i^a + \sum_{\substack{i<j \\ a<b}} C_{ij}^{ab} \Phi_{ij}^{ab} + \cdots , \tag{1.12}$$

where Φ_i^a and Φ_{ij}^{ab} are the singly and doubly excited determinants, and C_0, C_i^a, etc. are expansion coefficients that are determined with the variational method (i.e. the minimization of the CI energy). In CI calculations, the accuracy of results depends on both the level of excitation and the basis set that is used to construct excited determinants. If all the possible excitations could be taken into consideration and the complete basis set could be used, the resulting "full CI" would be the exact solution to the electronic Schrödinger equation. However, the computational cost of the full CI scales factorially with the number of electrons and basis functions. Therefore, truncated CI methods with a finite basis set are used in practice. In the CISD approximation (truncated at single and double excitations), for example, the determinant basis consists of the reference, Φ_0, single excitations, Φ_i^a, and double excitations, Φ_{ij}^{ab}, and it has a scaling of $\mathcal{O}(N^6)$. For methods beyond CISD, the computation can become very expensive.

While the full CI method is size-consistent (i.e. the energy of n isolated fragments is equal to n times the energy of one fragment) and size-extensive (i.e. the energy of a system scales linearly with the number of particles in the systems), the truncated CI methods are not. The size-consistency is particularly important for the proper description of dissociation energies, and the lack of this property in truncated CI methods has considerably reduced their applicability.[26] The CI theory has played an important role in the development of correlation methods, and was the most often used correlation method in the early days of quantum chemistry. Today, the MBPT and CC models, which are both size-consistent and size-extensive, have replaced the CI methods and become the preferred methods. Nevertheless, there is still continuous interest in the CI theory in the multireference context, whereas the multireference CC theory is more difficult.

1.3.2 Coupled-Cluster Methods

In the CC method,[27] the wave function is written as

$$\Psi_{\mathrm{cc}} = e^{\hat{T}}\Phi_0 = \left(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \cdots\right)\Phi_0 \tag{1.13}$$

with

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots, \tag{1.14}$$

where the exponentiated cluster operator is used to generate excited determinants. The n-electron cluster operator is defined as

$$\hat{T}_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\ldots ab\ldots}^{n} t_{ab\ldots}^{ij\ldots} a_a^\dagger b_b^\dagger \cdots a_j a_i \tag{1.15}$$

in the second quantization form, where a_i represents an annihilation operator, which removes an electron from the orbital ψ_i in the determinant when acting on the reference determinant, and $a_a^\dagger$ is a creation operator, which adds an electron to the orbital ψ_a in the determinant.

In the traditional CC theory, the electronic Schrödinger equation, $\hat{H}e^{\hat{T}}|\Phi_0\rangle = Ee^{\hat{T}}|\Phi_0\rangle$, is reformulated by multiplying $e^{-\hat{T}}$ on the left side:

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E|\Phi_0\rangle, \tag{1.16}$$

where $e^{-\hat{T}}\hat{H}e^{\hat{T}}$ is called the similarity-transformed Hamiltonian. This similarity-transformed Hamiltonian can be transformed into a linear combination of nested commutators of $\hat{H}$ and $\hat{T}$:

$$\begin{aligned} e^{-\hat{T}}\hat{H}e^{\hat{T}} = \hat{H} + [\hat{H},\hat{T}] + \frac{1}{2!}[[\hat{H},\hat{T}],\hat{T}] + \frac{1}{3!}[[[\hat{H},\hat{T}],\hat{T}],\hat{T}] \\ + \frac{1}{4!}[[[[\hat{H},\hat{T}],\hat{T}],\hat{T}],\hat{T}] + \cdots \end{aligned} \tag{1.17}$$

with the Baker-Campbell-Hausdorff (BCH) expansion, and it truncates itself at the four-fold commutator independent of the system or the cluster operator. Moreover, the CC amplitude equations are energy-independent with the similarity-transformed Hamiltonian:

$$\langle\Phi_{ij\ldots}^{ab\ldots}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0. \tag{1.18}$$

These amplitude equations are obtained by projecting the excited determinants on both sides of Eq. 1.16. Similarly, the energy equation is obtained by projecting the reference determinant on both sides:

$$\langle\Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E. \tag{1.19}$$

If all cluster operators are included, the resulting *full* CC energy is equivalent to the full CI value. However, the cluster operator is usually truncated at a finite rank in practice,

which leads to a hierarchy of coupled-cluster methods. In the coupled-cluster singles and doubles (CCSD) method, for example, $\hat{T} \equiv \hat{T}_1 + \hat{T}_2$. Although the CCSD method is relatively expensive (scales as $\mathcal{O}(N^6)$), it is often not accurate enough (i.e. above chemical accuracy, > 4.18 kJ/mol). Take the reaction energy of $2CH_2 \rightarrow C_2H_4$ (Figure 1.1) as an example: even with the large aug-cc-pVQZ basis set, the CCSD result has a error of 20.1 kJ/mol with respect to the experimental value. On the other hand, the CCSD(T) model, which includes a perturbative treatment of triples, can give results close to the experiment. For many properties, the CCSD(T) model has been shown to give consistent highly accurate results, and it is considered to be the "gold-standard of quantum chemistry." [28]

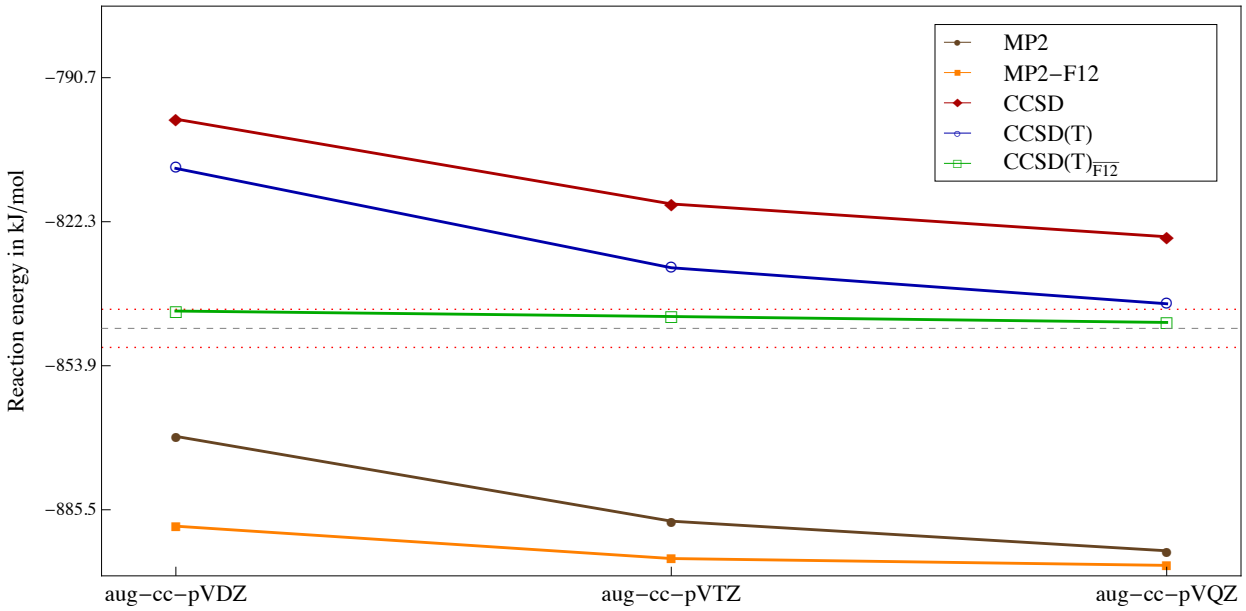

Figure 1.1: Calculated reaction energies (kJ/mol) of $2CH_2 \rightarrow C_2H_4$ with various correlation methods. The dashed line represents the experimentally derived value,[1] and the red lines denote chemical accuracy. Valence conventional and explicitly correlated calculations were performed, and thus the core correlation contribution from the reference[1] was added to the results.

1.3.3 Many-Body Perturbation Theory

Perturbation theory is one of the mathematical tools which has been used since early quantum chemistry to include descriptions of electron correlations in calculations. It provides a systematic procedure to obtain the correlation energy of a system. In MBPT,[22] the Hamiltonian is partitioned into the zeroth-order and a small perturbation:

$$\hat{H} = H^{(0)} + \lambda \hat{V}, \tag{1.20}$$

where the eigenfunctions and eigenvalues for the zeroth-order Hamiltonian are known, and λ is a perturbation-strength parameter and set to unity in the end. Furthermore, the exact energy and wave function can each be expressed as an infinite sum of corrections from different orders:

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots, \tag{1.21}$$

$$\Psi = \Psi^{(0)} + \lambda \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \cdots. \tag{1.22}$$

By inserting the above expressions into the electronic Schrödinger equation and then equating the same order of λ, the expressions for energies of different orders can be obtained:

$$E^{(0)} = \langle \Psi^{(0)} | H^{(0)} | \Psi^{(0)} \rangle, \tag{1.23}$$

$$E^{(1)} = \langle \Psi^{(0)} | \hat{V} | \Psi^{(0)} \rangle, \tag{1.24}$$

$$E^{(2)} = \langle \Psi^{(0)} | \hat{V} | \Psi^{(1)} \rangle, \tag{1.25}$$

$$\vdots$$

In the Møller-Plesset perturbation theory, the HF Hamiltonian is chosen as the zeroth-order Hamiltonian:

$$H^{(0)} = \sum_i \hat{F}_i, \tag{1.26}$$

and the perturbation is then given by

$$\hat{V} = \sum_{i<j} \frac{1}{r_{ij}} - \sum_i v_{\mathrm{HF}}(i). \tag{1.27}$$

As a result, the first-order wave function is the HF wave function, and the sum of the zeroth- and first-order energy is the HF energy. The first correlation correction to the HF energy appears in the second-order Møller-Plesset method (MP2):

$$E^{(2)} = \frac{1}{4} \sum_{ijab} \frac{|\langle \Phi_0 | \hat{V} | \Phi_{ij}^{ab} \rangle|^2}{\varepsilon_i + \varepsilon_i - \varepsilon_a - \varepsilon_b}. \tag{1.28}$$

The MP2 method recovers about 80 to 90% of the correlation energy. With higher order corrections, such as MP3 and MP4 corrections, the correlation energy can be described more accurately. In some cases, however, MPn energies have been shown to diverge as n increases,[1] and MP2 is the most used method in the series. Although the MP2 method is less accurate than the CCSD or CISD methods, it has a lower scaling as $\mathcal{O}(N^5)$. Moreover, it is both size-consistent and size-extensive. Thus MP2 represents a good method for studies of moderate-size systems.

1.4 Density Functional Theory

An alternative approach to standard correlation methods is the density functional theory (DFT),[29] in which correlation effects are included implicitly. In this approach, the explicit use of the many-electron wave function is avoided. According to the Hohenberg-Kohn theorem,[30] there is a unique functional relationship between the electron density and the Hamiltonian for a given system, and all the properties of the system can be parametrized through the electron density. Therefore, the energy can be expanded as:

$$E_0 = E_0[\rho_0] = \overline{T}[\rho_0] + \overline{N}_e[\rho_0] + \overline{V}_{ee}[\rho_0], \tag{1.29}$$

where E_0 is the electronic energy of the ground state, ρ_0 is the corresponding electron density that must be modeled in practice, and $\overline{T}[\rho_0]$, $\overline{N}_e[\rho_0]$, and $\overline{V}_{ee}[\rho_0]$ correspond to the electronic kinetic energy, electron-nucleus attraction, and electron-electron repulsion, respectively. Practical implementation of DFT uses the Kohn-Sham (KS) formulation,[31] which is based on a single Slater-determinant wave function. The difference between the KS-DFT and HF methods is that KS-DFT accounts for the electron correlation and exchange effects via an exchange-correlation (XC) functional of the density. If the exact functional could be obtained, KS-DFT methods would give the exact solution to the electronic Schrödinger equation. In practice, however, the functional needs to be modeled. Modern KS-DFT methods usually use hybrid density functionals (e.g., B3LYP), which are linear combinations of the HF exchange, DFT exchange, and DFT correlation functionals.

Compared with standard correlation methods, DFT methods have relatively low computational cost ($\mathcal{O}(N^4)$ as opposed to $\mathcal{O}(N^5)$ to $\mathcal{O}(N^7)$ for standard correlation methods, where N is the size of the system), but give relatively accurate results (comparable to MP2, or better). Therefore, DFT has gained widespread popularity in chemistry over the last two decades. However, DFT recovers electron correlation effects in an empirical way,[32] which means there is no systematic way to improve the accuracy of DFT calculations. Thus, DFT is considered a "non-convergent" method, and wave function methods are the only means to reliably approach the exact solution of the electronic Schrödinger equation.

1.5 Slow Convergence Problem of Correlation Energy

As we mentioned earlier, the accuracy of correlation energies in standard correlation methods depends on both the level of excitation and the basis set which is used to construct excited determinants. Therefore, the choices for the level of excitation and the basis set need to be balanced. There is little gain in accuracy to include higher excitations when the basis set errors are larger than or close to errors from neglecting higher excitations. Unfortunately, the decay of errors in correlation energies is inversely proportional to the size of the basis set.[33] This means the size of the basis set needs to be increased by a factor of ten for an order of magnitude improvement in the accuracy. Moreover, this slow convergence problem is coupled

with a steep increase of the computational cost as the basis size increases. Since four-index two-electron integrals are involved in all standard correlation methods, the computational cost of calculations with those methods scales as, at least, $\mathcal{O}(K^4)$ (K is the number of basis functions), which results in 10,000 fold increase in the computational time for one digit reduction in the basis set error.

The sole use of Slater determinants in standard correlation methods is the reason for the slow convergence problem of correlation energies with respect to the basis set size. This is because these Slater determinants are constructed with products of one-electron orbitals, but the Coulomb hole that appears at short inter-electronic distances cannot be efficiently described by products of one-electron orbitals alone.[34] This issue can be simply illustrated with the helium atom. The nonrelativistic electronic Hamiltonian for the helium atom is written as

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}. \tag{1.30}$$

As the two electrons approach each other ($r_{12} \to 0$), the Coulomb potential will approach infinity, and the Hamiltonian will become singular. To satisfy the Schrödinger equation, $\hat{H}\Psi = E\Psi$, the singularity must be balanced by an opposite infinity in the kinetic energy. This can be demonstrated by the Hamiltonian for the helium ground state written in the electron-nucleus distances (r_1 and r_2) and the electron-electron distances (r_{12}):[35]

$$\begin{aligned}\hat{H} &= -\frac{1}{2}\sum_{i=1}^{2}\left(\frac{\partial^2}{\partial r_i^2} + \frac{2}{r_i}\frac{\partial}{\partial r_i} + \frac{2Z}{r_i}\right) - \left(\frac{\partial^2}{\partial r_{12}^2} + \frac{2}{r_{12}}\frac{\partial}{\partial r_{12}} - \frac{1}{r_{12}}\right) \\ &\quad - \left(\frac{\mathbf{r_1}}{r_1}\cdot\frac{\mathbf{r_{12}}}{r_{12}}\frac{\partial}{\partial r_1} + \frac{\mathbf{r_2}}{r_2}\cdot\frac{\mathbf{r_{21}}}{r_{21}}\frac{\partial}{\partial r_2}\right)\frac{\partial}{\partial r_{12}},\end{aligned} \tag{1.31}$$

where the singularity at the electron coalescence can be cancelled by the kinetic energy term, $2/r_{12}(\partial/\partial r_{12})$. This requirement leads to the Coulomb cusp condition:[36]

$$\left.\frac{\partial\Psi}{\partial r_{12}}\right|_{r_{12}=0} = \frac{1}{2}\Psi(r_{12}=0), \tag{1.32}$$

which describes the behavior of the wave function when the electrons coalesce. The Coulomb cusp condition suggests that the wave function for the singlet states becomes linear in r_{12} when the two electrons are close,

$$\Psi(\mathbf{r_1}, \mathbf{r_2}) = \Psi(r_{12}=0)\big(1 + \frac{1}{2}r_{12} + \cdots\big). \tag{1.33}$$

For triplet states, there is no infinity in the Schrödinger equation as the wave function becomes zero at the points of the electron coalescence. Yet the wave function for triplet states of the helium atom also has a discontinuity but in the second derivative:

$$\left.\frac{\partial^2\Psi}{\partial r_{12}^2}\right|_{r_{12}=0} = \frac{1}{2}\left.\frac{\partial\Psi}{\partial r_{12}}\right|_{r_{12}=0}. \tag{1.34}$$

This condition indicates that the behavior of the triplet helium wave function near the electron coalescence has the following form:

$$\Psi(\mathbf{r_1}, \mathbf{r_2}) \;=\; \mathbf{r_{12}} \cdot \mathbf{w}\big(1 + \frac{1}{4} r_{12} + \cdots\big) \tag{1.35}$$

where $\mathbf{w}$ is the vector $(\partial\Psi/\partial x_{12}, \partial\Psi/\partial y_{12}, \partial\Psi/\partial z_{12})$ evaluated at the electron coalescence.

These conditions are rigorously observed by the exact wave functions, but not CI-like wave functions (Eq. 1.12). As a consequence, very high angular momenta must be included in the basis set to describe the short-range two-electron correlation.[2] A more efficient way to describe electron correlations is to include r_{12} explicitly in wave function expansions, and thus the basis set convergence of correlation energies can be improved. Methods where the approximate wave functions depend explicitly on r_{12} are called explicitly correlated methods. In Figure 1.1, for example, the CCSD(T) reaction energy converges very slowly to the experimental value, and only with the large aug-cc-pVQZ basis set does it approach the experimental value. However, the explicitly correlated method, CCSD(T)$_{\overline{\text{F12}}}$, gives a reaction energy near the experimental value even at the double-ζ level.

Chapter 2

Overview of Early Explicitly Correlated Methods

The use of inter-electronic distances in wave functions dates back to the beginning of quantum mechanics in the late 1920s.[35] Early explicitly correlated methods are only applicable to very small systems; their extensions to many-electron systems are very difficult. The major hurdle for the applications of explicitly correlated methods to systems with more than two electrons is the evaluation of numerous and expensive many-electron integrals, which result from the use of functions that depend explicitly on inter-electronic distances. Over the years, various strategies have been developed to overcome this technical challenge. These efforts include the use of Gaussian geminals,[37,38] similarity-transformed Hamiltonian,[39,40] and, more recently, the resolution of identity[9] in the R12 methods.

2.1 Methods for Two-Electron Systems

2.1.1 Hylleraas Expansion

In 1929, Hylleraas realized that the slow convergence problem for the helium atom can be overcome by including the r_{12} term explicitly in the wave function.[35] The wave function he employed has the following form:

$$\Psi = e^{-\zeta s}(c_1 + c_2 u + c_3 t^2), \tag{2.1}$$

where the coordinates: $s = r_1 + r_2$, $t = r_1 - r_2$, and $u = r_{12}$ are used, and the exponential coefficient, ζ, and the expansion coefficients, c_i, can be determined variationally. By using this three-term wave function with $\zeta = 1.82$, he was able to obtain the ground-state energy of helium as -2.90243 E_h. This result differs from the exact Born-Oppenheimer nonrelativistic value by 1.3 mE_h (3.4 kJ/mol), which was a dramatic breakthrough at that time. It should

be noted that the success of the Hylleraas wave function is not merely due to the inclusion of the linear r_{12} term. The inclusion of both r_{12} and $(r_1 - r_2)^2$ are necessary, and the reoptimization of the orbital shape, via the adjustment of exponent ζ, is also important.[2]

A general form of the Hylleraas expansion for a ^1S state of a two-electron atom can be written as:

$$\Psi = e^{-\zeta s} \sum_{i=1}^{M} c_i s^{l_i} t^{2m_i} u^{n_i}, \tag{2.2}$$

where l_i, m_i, and n_i are non-negative integers. These Hylleraas expansions have been shown to converge much more rapidly than CI wave functions. Figure 2.1 presents the errors of helium ground-state energies computed with a Hylleraas expansion and a conventional CI wavefunction.[2] It is clear that the Hylleraas expansion outperforms the CI expansion, and approaches spectroscopic accuracy (1 cm^{-1}) very rapidly.

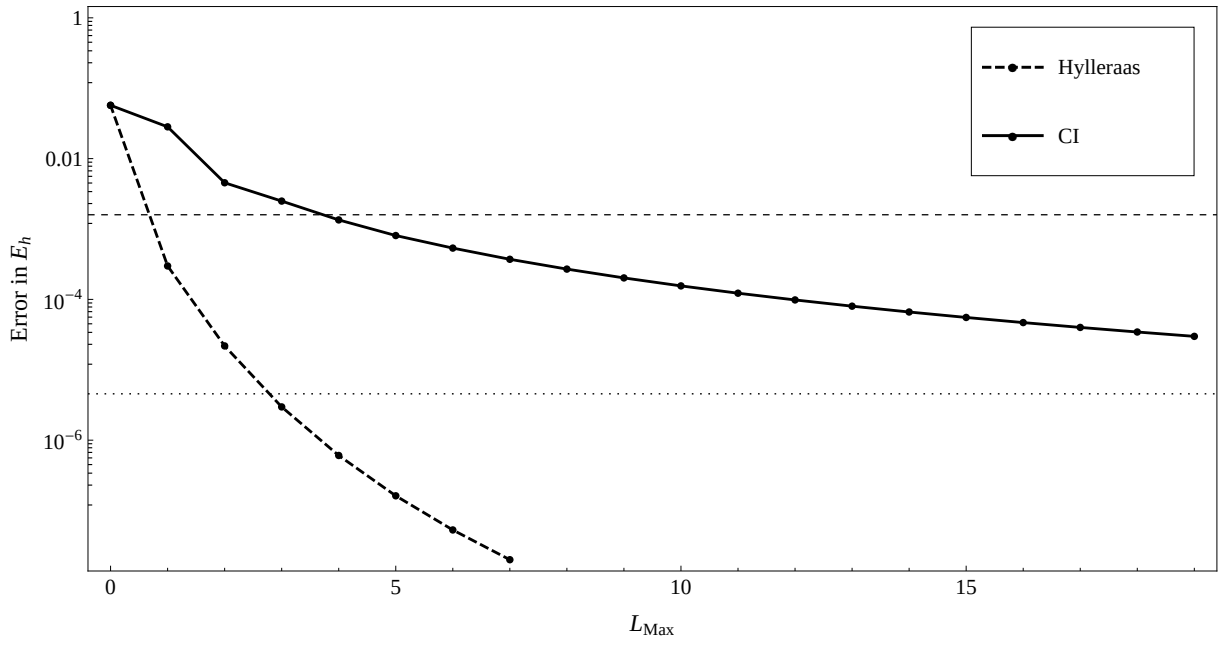

Figure 2.1: The errors of the helium ground-state energy computed with the Hylleraas expansion and the conventional CI wavefunction (reproduced with data from E. F. Valeev, original figure published in *Chem. Rev.*, 2012, **112**, 75.). The dashed line denotes chemical accuracy, and the dotted line represents spectroscopic accuracy. The Hylleraas wave function includes all terms $k + l + m \leq 2L_{\text{max}}$, and has an orbital exponent fixed at 1.8149. The CI expansion was computed with a basis sets of modified hydrogenic orbitals with angular momentum up to L_{max} and principal quantum number up to $L_{\text{max}} + 1$ (see Ref. 2 for more details).

The Hylleraas-type wave functions have been subsequently applied to helium-like, and even lithium-like, atoms.[9,41] It is possible to generalize the Hylleraas approach to many-electron systems. For example, the Hylleraas-CI method has been applied to systems with up to ten electrons.[42] However, numerous many-electron integrals appear in the formalism due to the inter-electronic distances, and their evaluation quickly becomes very expensive and complicated as the number of electrons increases, which severely limits the application of the Hylleraas method.

2.1.2 James-Coolidge Wave Function

In 1933, James and Coolidge extended the Hylleraas approach to the H_2 molecule,[43] which is the first molecular system studied with explicitly correlated wave functions. Their spin-free singlet wave function for the H_2 molecule is written as:

$$\Psi_{\mathrm{JC}} = \sum_i C_i\big(\Psi_i(1,2) + \Psi_i(2,1)\big) \tag{2.3}$$

with

$$\Psi_i(1,2) = e^{-\alpha(\xi_1+\xi_2)}\xi_1^{n_i}\eta_1^{k_i}\xi_2^{m_i}\eta_2^{l_i}r_{12}^{\mu_i}. \tag{2.4}$$

Here,

$$\xi_i = \frac{(r_{iA}+r_{iB})}{R}, \tag{2.5}$$

$$\eta_i = \frac{(r_{iA}-r_{iB})}{R} \tag{2.6}$$

are the elliptic coordinates, A and B represent the hydrogen nuclei, R is the internuclear distance, k_i, m_i, l_i and μ_i are non-negative integers, and α is a variational parameter. With a thirteen-term James-Coolidge (JC) expansion, the energy calculated for the ground state of H_2 has an error of 0.917 $\mathrm{mE_h}$ with respect to the best variational value (-1.174476 $\mathrm{E_h}$). Despite this, the JC wave function fails to give a correct description for the dissociation behavior of H_2.

In 1965, Kołos and Wolniewicz improved the JC wave function with expansions in terms of the following basis functions:[44]

$$\Psi_i(1,2) = e^{-\alpha\xi_1-\bar{\alpha}\xi_2}\xi_1^{n_i}\eta_1^{k_i}\xi_2^{m_i}\eta_2^{l_i}r_{12}^{\mu_i}\left(e^{\beta\eta_1+\bar{\beta}\eta_2} + (-1)^{k_i+l_i}e^{-\beta\eta_1-\bar{\beta}\eta_2}\right), \tag{2.7}$$

where $\bar{\alpha}$ and $\bar{\beta}$ are also nonlinear variational parameters as α and β. With such a basis, the Kołos-Wolniewicz (KW) function correctly describes the dissociation of H_2. In principle, it is possible to extend the KW function to more than two electrons systems. However, the evaluation of many-electron integrals that result from inter-electronic distances is still a limiting factor for its application to many-electron systems.

2.2 Methods for Many-Electron Systems

2.2.1 Gaussian Geminal Methods

In Gaussian geminal methods, the use of explicitly correlated Gaussian functions allows the many-electron integrals to be evaluated analytically. In the two-electron case, these functions are referred to as Gaussian geminals, which can be in general expressed as

$$g_k(\mathbf{r}_1, \mathbf{r}_2) = \exp(-\alpha_k|\mathbf{r}_1 - \mathbf{A}_k|^2 - \beta_k|\mathbf{r}_2 - \mathbf{B}_k|^2 - \gamma_k r_{12}^2) \tag{2.8}$$

where r_{12} is the inter-electronic distance, and the exponents (α_k, β_k, and γ_k) and the centers ($\mathbf{A}_k$ and $\mathbf{B}_k$) are all variational parameters. A single Gaussian geminal does not satisfy the cusp condition, but a linear combination of these functions with optimized parameters can effectively describe this behavior. This idea was first proposed by Boys[37] and Singer[38] in 1960. Since then, it has been used successfully for many applications. For example, the calculations with Gaussian geminal methods on H_2 and He_2 gave highly accurate predictions for the energies.[45,46] However, the complexity of these methods grows rapidly with the number of electrons in the system. Thus early Gaussian geminal methods are only feasible for systems with four or fewer electrons.

For larger systems, modern Gaussian geminal methods[47–51] utilize geminal functions in the framework of MBPT or CC methods, in which the HF determinant is the starting point and geminal functions are only included in the form of pair functions. The coupled-cluster doubles (CCD) wave function, for example, is

$$\begin{aligned} \Psi_{\mathrm{CCD}} &= e^{\hat{T}_2}\Phi_0 \\ &= \Psi_0 + \sum_{\substack{i<j \\ a<b}} t_{ab}^{ij}\Psi_{ij}^{ab} + \frac{1}{2}\sum_{\substack{i<j \\ a<b}}\sum_{\substack{k<l \\ c<d}} t_{ab}^{ij}t_{cd}^{kl}\Psi_{ijkl}^{abcd} + \cdots \end{aligned} \tag{2.9}$$

where t_{ab}^{ij} are undetermined double excitation amplitudes, and $\hat{T}_2$ is the two-electron cluster operator used to generate doubly substituted (or excited) determinants. The two-electron cluster operator can be written as

$$\hat{T}_2 = \sum_{i<j} |\tau_{ij}\rangle\langle ij| \tag{2.10}$$

with pair functions:

$$|\tau_{ij}\rangle = \sum_{a<b} t_{ab}^{ij}|ab\rangle, \tag{2.11}$$

where i and j represent the occupied orbitals while a and b represent the virtual orbitals. In Gaussian geminal methods, pair functions are expanded in terms of Gaussian geminals:

$$|\hat{\tau}_{ij}\rangle = \hat{Q}_{12}\sum_k t_k^{ij}|g_k^{ij}(\mathbf{r}_1, \mathbf{r}_2)\rangle. \tag{2.12}$$

In the new pair functions, t_k^{ij} are geminal amplitudes to be determined, and $\hat{Q}_{12}$ is a projection operator that ensures pair functions are strongly orthogonal to products of occupied orbitals.

When Gaussian geminals are employed within the MBPT and CC frameworks, up to five-electron integrals can appear. The evaluation of these integrals becomes challenging for larger systems. In an attempt to reduce the complexity of the methods, approximations such as weak orthogonality (WO), super weak orthogonality (SWO), and SWO plus projection (SWOP) were introduced.[47,48] With these approximations, calculations with Gaussian geminals have been successfully applied to systems with up to ten electrons.[49,50]

2.2.2 Exponentially Correlated Gaussian Methods

Explicitly correlated Gaussian functions are also used in exponentially correlated Gaussian (ECG) methods.[52–54] The approximate wave function in these methods is expanded with a basis set of n-electron functions, which correlate the motion of all pairs of electrons. The spatial part of an N-electron basis function takes the following form:

$$\Psi_k(\mathbf{r}_1, \cdots, \mathbf{r}_n) = \exp\left(-\sum_{i=1}^{n} \alpha_i^k |\mathbf{r}_i - \mathbf{C}_i^k|^2 - \sum_{i<j=1}^{n} \beta_{ij}^k |\mathbf{r}_i - \mathbf{r}_j|^2\right). \tag{2.13}$$

Due to the Gaussian form, n-electron integrals in ECG methods can be evaluated analytically.[52] As encountered in Gaussian geminal methods, however, the computational cost of n-electron integrations is very expensive. Moreover, a large number of nonlinear parameters need to be optimized in the process of energy minimization due to the use of n-electron basis functions. As a result, ECG methods have only been applied to systems with up to four electrons. For small molecules, ECG methods are an effective approach for highly accurate calculations. For example, a 2400-term ECG expansion has been used to compute the ground-state energy of LiH, and the resulting error is $15\mu E_h$.[55]

2.2.3 Transcorrelated Methods

In the transcorrelated method of Boys and Handy,[39,40] a similarity-transformed Hamiltonian,

$$\bar{H}_G = e^{-\hat{G}} \hat{H} e^{\hat{G}} \tag{2.14}$$

with

$$\hat{G} = \sum_{i<j} f(\mathbf{r}_i, \mathbf{r}_j), \tag{2.15}$$

is used. The similarity-transformed Hamiltonian contains only one-, two-, and three-electron operators, as the Hausdorff expansion,

$$\bar{H}_G = \hat{H} + [\hat{H}, \hat{G}] + \frac{1}{2}[[\hat{H}, \hat{G}], \hat{G}] + \cdots, \tag{2.16}$$

truncates at the quadratic (third) term. Therefore, only two- and three-electron integrals are involved in this formulation. In Handy's study, the correlation function is comprised of Gaussian geminals and Gaussian atomic functions:

$$f(\mathbf{r}_i, \mathbf{r}_j) = \sum_{\kappa} c_\kappa g_\kappa(\mathbf{r}_i, \mathbf{r}_j) + \sum_{\lambda} d_\lambda \big(g_\lambda(\mathbf{r}_i) + g_\lambda(\mathbf{r}_j)\big), \tag{2.17}$$

which allows the many-electron integrals involved to be computed analytically. With a small expansion of $f(\mathbf{r}_i, \mathbf{r}_j)$, highly accurate results can be obtained for small systems (e.g., atoms and LiH). However, the transformed Hamiltonian is non-Hermitian, which means the energies calculated with the transcorrelated method are not variational. As a result, a more flexible, one-electron basis set is required to obtain stable results compared to other explicitly correlated methods.

Following the studies of Boys and Handy, Ten-no[56] developed the transcorrelated method with a frozen Gaussian geminals. In this approach, the Coulomb repulsion is canceled at short inter-electronic distances. It is possible to apply this transcorrelated method to large-scale systems, but the quality of the results was not encouraging. Recently, Luo et al.[57] proposed a new variational transcorrelated method, and they demonstrated that it could be an efficient method to deal with electron correlations.

Chapter 3

Modern R12/F12 Methods

In 1985, Kutzelnigg[9] proposed a new explicitly correlated method: the R12 method, wherein he also introduced the resolution of the identity (RI) to solve the many-electron integral problem. In this approach, conventional wave functions are augmented with a reference determinant multiplied by linear r_{ij} factors, which can better describe the wave function in the region near the electron coalescence. This formalism is much simpler than those of many older explicitly correlated methods. Take the helium atom as an example: its R12 wave function can be written as a linear combination of the reference function (Φ), conventional correlation term(χ), and explicitly correlated term:

$$\Psi(\mathbf{r_1}, \mathbf{r_2}) \quad = \quad \Phi(\mathbf{r_1}, \mathbf{r_2}) + \chi(\mathbf{r_1}, \mathbf{r_2}) + cr_{12}\Phi(\mathbf{r_1}, \mathbf{r_2}). \tag{3.1}$$

Since the R12 wave function is merely a linear combination of Slater determinants augmented by a few explicitly correlated terms, it can, in principle, be implemented with any conventional method (e.g., MBPT or CC methods). The R12 approach was first implemented within the MP2 framework.[58] Many of the important developments of the R12 technology have also been made on the MP2 level. It has been shown that the R12 methods are very effective in reducing basis set errors of MP2 correlation energies. To approach the exact solution of the electronic Schrödinger equation, it is important to not only reduce the basis set error, but also incorporate high-level correlation effects. As a result, the R12 approach was later extended to coupled-cluster R12 (CC-R12) methods[59,60] and multi-reference configuration interaction R12 (MRCI-R12) methods.[61,62]

3.1 MP2-R12 Method

3.1.1 The Formalism

In the MP2 theory, the Hylleraas functional can be used to obtain the second-order energy:

$$H^{(2)} = 2\langle\Psi^{(1)}|\hat{V}|\Psi^{(0)}\rangle + \langle\Psi^{(1)}|\hat{H}^{(0)} - E^{(0)}|\Psi^{(1)}\rangle \geq E^{(2)}. \tag{3.2}$$

The first-order wave function in the Hylleraas functional is comprised of doubly excited determinants, and can be written as:

$$\Psi^{(1)}_{\text{MP}} = \frac{1}{4} t^{ij}_{ab} \tilde{a}^{ab}_{ij} |\Psi^{(0)}\rangle \tag{3.3}$$

with $\tilde{a}^{ab}_{ij} = a^\dagger_a a^\dagger_b a_j a_i$ in the second quantization form, where the Einstein summation convention is also used (i.e. summation is implied over repeated indices). In the MP2-R12 method,[2,63] the first-order wave function includes additional explicitly correlated (geminal) terms:

$$\Psi^{(1)}_{\text{MP-R12}} = \Psi^{(1)}_{\text{MP}} + \frac{1}{8} t^{ij}_{kl} \bar{R}^{kl}_{\alpha\beta} \tilde{a}^{\alpha\beta}_{ij} |\Psi^{(0)}\rangle, \tag{3.4}$$

where t^{ij}_{kl} are undetermined geminal amplitudes, and α, β, ... are indices representing the virtual orbitals in the complete basis. The antisymmetrized integral $\bar{R}^{ij}_{\alpha\beta}$ represents the tensor element of the geminal correlation factor $f(r_{12})$:

$$R^{ij}_{\alpha\beta} = \langle\alpha\beta|\hat{Q}_{12} f(r_{12})|ij\rangle. \tag{3.5}$$

The projector $\hat{Q}_{12}$ ensures that the geminal functions are strongly orthogonal to the reference and the standard doubly excited determinants:

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1 \hat{V}_2, \tag{3.6}$$

where $\hat{O}$ and $\hat{V}$ are the projectors on the occupied and virtual orbitals. The resulting MP2-R12 correlation energy is a sum of the conventional MP2 energy and the R12 correction from geminal terms:

$$E^{(2)}_{\text{MP2-R12}} = E^{(2)}_{\text{MP2}} + E^{(2)}_{\text{R12}}. \tag{3.7}$$

The R12 correction is computed as

$$E^{(2)}_{\text{R12}} = -\frac{1}{4} t^{ij}_{kl} \tilde{V}^{kl}_{ij} \tag{3.8}$$

with

$$t^{ij}_{kl} = \frac{1}{2} \tilde{V}^{ij}_{mn} \big(\tilde{B}^{-1}_{(ij)}\big)^{mn}_{kl}, \tag{3.9}$$

where $\tilde{V}$ and $\tilde{B}$ are intermediates of the R12 theory. $\tilde{V}$ is the geminal tensor of the perturbation, and its element has the following form:

$$\tilde{V}_{ij}^{kl} = V_{ij}^{kl} + \frac{1}{2} C_{ab}^{kl} t_{ij}^{ab} \tag{3.10}$$

with

$$V_{ij}^{kl} = \frac{1}{2} \bar{g}_{ij}^{\alpha\beta} \bar{R}_{\alpha\beta}^{kl}, \tag{3.11}$$

$$C_{ab}^{kl} = \frac{1}{2} (f_a^\alpha \bar{R}_{\alpha b}^{kl} + f_b^\alpha \bar{R}_{a\alpha}^{kl}). \tag{3.12}$$

Here, f_q^p represents a Fock matrix element, $f_p^q = \langle p|\hat{F}|q\rangle$, and $\bar{g}_{rs}^{pq}$ represents an antisymmetrized Coulomb integral, $\bar{g}_{pq}^{rs} = \langle pq|r_{12}^{-1}|rs\rangle - \langle pq|r_{12}^{-1}|sr\rangle$. $\tilde{B}$ is the geminal tensor of the zero-order Hamiltonian. If we use canonical HF orbitals, the tensor element of $\tilde{B}$ can be expressed as

$$(\tilde{B}_{(ij)})_{kl}^{mn} = B_{kl}^{mn} - (f_i^i + f_j^j) X_{kl}^{mn} - \frac{1}{2} \frac{C_{kl}^{ab} C_{ab}^{mn}}{f_a^a + f_b^b - f_i^i - f_j^j}, \tag{3.13}$$

where

$$B_{kl}^{mn} = \bar{R}_{\alpha\gamma}^{mn} f_\beta^\alpha \bar{R}_{kl}^{\beta\gamma}, \tag{3.14}$$

$$X_{kl}^{mn} = \frac{1}{2} \bar{R}_{\alpha\beta}^{mn} \bar{R}_{kl}^{\alpha\beta}. \tag{3.15}$$

3.1.2 Evaluation of Many-Electron Integrals in MP2-R12 Theory

Standard Approximation

Similar to other explicitly correlated methods, many-electron integrals appear in the MP2-R12 method due to the inclusion of explicitly correlated terms. In the intermediates V and X, up to three-electron integrals are involved, whereas the computation of the B intermediate requires up to four-electron integrals. To avoid computing these many-electron integrals, Kutzelnigg and Klopper proposed a set of approximations known as Standard Approximation (SA).[64] In the SA, the Brillouin condition,

$$f_a^i = f_i^a = 0, \tag{3.16}$$

is assumed. Moreover, the generalized Brillouin condition (GBC),

$$f_{a'}^i = f_i^{a'} = 0, \ \forall i, a', \tag{3.17}$$

and extended Brillouin condition (EBC),

$$f_{b'}^a = f_a^{b'} = 0, \ \forall a, b', \tag{3.18}$$

are also adopted, where a', b', ... denote indices for the virtual orbitals in the complete basis that do not belong to the orbital basis set (OBS), e.g., the Hartree-Fock basis. More importantly, the resolution of the identity (RI),

$$1 = \sum_{\kappa} |\kappa\rangle\langle\kappa| \approx \sum_{p'} |p'\rangle\langle p'| \tag{3.19}$$

is introduced in the SA, where $\{\kappa\}$ is a complete basis set, and $\{p'\}$ is the RI basis set that approximates the complete basis set. The use of the RI factorizes many-electron integrals into products of two-electron integrals. In the following, we will discuss how the RI facilitates the evaluation of V, X, and B intermediates as well as different strategies to approximate the RI.

Evaluation of V and X Intermediates

Tensor element V_{ij}^{kl} can be rewritten as

$$\begin{aligned} V_{ij}^{kl} &= \langle ij|\frac{1}{r_{12}}\hat{Q}_{12}f_{12}|kl\rangle \\ &= \langle ij|\frac{1}{r_{12}}\left((1-\hat{O}_1)(1-\hat{O}_2) - \hat{V}_1\hat{V}_2\right)f_{12}|kl\rangle \\ &= \langle ij|\frac{f_{12}}{r_{12}}|kl\rangle + \sum_{mn}\langle ij|\frac{1}{r_{12}}|mn\rangle\langle mn|f_{12}|kl\rangle - \sum_{ab}\langle ij|\frac{1}{r_{12}}|ab\rangle\langle ab|f_{12}|kl\rangle \\ &- \langle ij|\frac{1}{r_{12}}\hat{O}_1 f_{12}|kl\rangle - \langle ij|\frac{1}{r_{12}}\hat{O}_2 f_{12}|kl\rangle. \end{aligned} \tag{3.20}$$

In this expression, the first three terms are two-electron integrals, which can be evaluated analytically if Gaussian basis sets are used. The last two terms, however, are three-electron integrals:

$$\langle ij|\frac{1}{r_{12}}\hat{O}_1 f_{12}|kl\rangle = \sum_{m}\langle ijm|\frac{1}{r_{12}}f_{23}|mlk\rangle, \tag{3.21}$$

$$\langle ij|\frac{1}{r_{12}}\hat{O}_2 f_{12}|kl\rangle = \sum_{m}\langle ijm|\frac{1}{r_{12}}f_{23}|kml\rangle. \tag{3.22}$$

Nevertheless, they can be reduced to products of two-electron integrals with the RI approximation. For example,

$$\begin{aligned} \langle ijm|\frac{1}{r_{12}}f_{23}|mlk\rangle &= \sum_{\kappa}\langle ijm|\frac{1}{r_{12}}|\kappa\rangle\langle\kappa|f_{23}|mlk\rangle \\ &= \sum_{\kappa}\langle ij|\frac{1}{r_{12}}|m\kappa\rangle\langle m\kappa|f_{23}|mlk\rangle \\ &\approx \sum_{p'}\langle ij|\frac{1}{r_{12}}|mp'\rangle\langle mp'|f_{23}|mlk\rangle, \end{aligned} \tag{3.23}$$

where the complete basis set $\{\kappa\}$ is approximated by an auxiliary basis set (ABS) $\{p'\}$. In early R12 work, the orbital (i.e. Hartree-Fock) basis set, which is used for the expansion of the wave function, was employed to approximate the RI.[64] As a result, a very large orbital basis set was required for high accuracy, which limited the applications of those methods to atoms and small molecules. This situation was improved by Klopper and Samson,[65] who introduced a separate, auxiliary basis set (ABS) to approximate the RI while a relatively small basis set was used for the wave function. Later, Valeev proposed a modified scheme, complementary auxiliary basis set (CABS) approach[66] that has significantly smaller RI errors compared to the ABS approach. In the CABS approach, the $\hat{Q}_{12}$ projector is rewritten as

$$\hat{Q}_{12} = 1 - \hat{O}_1\hat{V}_2' - \hat{V}_1'\hat{O}_2 - \hat{P}_1\hat{P}_2, \tag{3.24}$$

where $\hat{V}'$ is the projector on the orbitals from the complete basis set that are complementary (i.e. orthogonal) to the OBS, and $\hat{P}$ represents the projector on the OBS. The $\hat{V}'$ projector is approximated with a sum over the CABS:

$$\hat{V}' \approx \sum_{a'} |a'\rangle\langle a'|, \tag{3.25}$$

where a' represent orbitals in the CABS space and are constructed by finding the null space of the overlap matrix between the auxiliary and orbital basis sets using singular value decomposition.[66] Thus, in the CABS approach the final expression for the V intermediate can be written as

$$V_{ij}^{kl} = \left(\frac{\overline{f_{12}}}{r_{12}}\right)_{ij}^{kl} - \frac{1}{2}\bar{r}_{pq}^{kl}\bar{g}_{ij}^{pq} - \bar{r}_{ma'}^{kl}\bar{g}_{ij}^{ma'}, \tag{3.26}$$

where $\bar{r}_{rs}^{pq}$ represent the antisymmetrized integrals of the bare correlation factor, $f(r_{12})$, and $\bar{g}_{rs}^{pq}$ are the antisymmetrized Coulomb integrals. Similarly, the intermediate X can be evaluated with the following expression:

$$X_{ij}^{kl} = \left(\overline{f_{12}^2}\right)_{ij}^{kl} - \frac{1}{2}\bar{r}_{pq}^{kl}\bar{r}_{ij}^{pq} - \bar{r}_{ma'}^{kl}\bar{r}_{ij}^{ma'}. \tag{3.27}$$

In addition to the standard Coulomb integrals, Eq. 3.26 and 3.27 also involve two-electron integrals over $f(r_{12})$, $f(r_{12})/r_{12}$, and $(f(r_{12}))^2$, and they can be evaluated at similar cost to that of the Coulomb integrals. Compared with the intermediates V and X, the evaluation of the intermediate C is much simpler since it contains only one- and two-electron integrals, and its formula in the CABS approach is written as

$$C_{ab}^{kl} = \frac{1}{2}(f_a^{a'}\bar{r}_{a'b}^{kl} + f_b^{a'}\bar{r}_{aa'}^{kl}). \tag{3.28}$$

Evaluation of Intermediate B

In the MP2-R12 theory, B, the geminal tensor element on the Fock operator, is the most complex intermediate. Its tensor elements can be written as

$$B_{kl}^{ij} = \langle ij|f_{12}\hat{Q}_{12}\big(\hat{F}_1 + \hat{F}_2\big)\hat{Q}_{12}f_{12}|kl\rangle. \tag{3.29}$$

In the early MP2-R12 approach, its derivation utilized the GBC assumption and a tricky manipulation of commutators:[64,65]

$$B_{kl}^{ij} \approx \frac{1}{2}\langle ij|f_{12}\hat{Q}_{12}[\hat{F}_{12},\hat{Q}_{12}f_{12}]|kl\rangle + \frac{1}{2}\langle ij|[f_{12}\hat{Q}_{12},\hat{F}_{12}]\hat{Q}_{12}f_{12}]|kl\rangle, \tag{3.30}$$

where $\hat{F}_{12} = \hat{F}_1 + \hat{F}_2$. The commutator $[\hat{F}_{12},\hat{Q}_{12}f_{12}]$ in the expression leads to integrals over $[\hat{F}_{12},f_{12}]$, which can be decomposed into:

$$[\hat{F}_{12},f_{12}] = [\hat{T}_{12},f_{12}] - [\hat{K}_{12},f_{12}], \tag{3.31}$$

Here, only the kinetic energy and exchange operators ($\hat{T}_{12}$ and $\hat{K}_{12}$) are involved as the nuclear potential and the Coulomb operator commute with f_{12}. While the integrals involving the kinetic energy commutator can be evaluated analytically, the integrals over the exchange commutator require double insertions of the RI, which leads to a quadratic scaling of the integral evaluation with respect to the size of the RI basis. In the standard approximation A,[63] $[f_{12},\hat{K}_{12}]$ is neglected, which leads to a much simpler formula and lower computational cost. In the standard approximation B, the integrals over $[f_{12},\hat{K}_{12}]$ are computed rigorously, and thus it is more computational demanding than approximation A. On the other hand, it has been shown that the MP2-R12/B method converges faster to the basis set limit than the MP2-R12/A method ($\mathcal{O}[(L+1)^{-7}]$ vs. $\mathcal{O}[(L+1)^{-5}]$), and the latter overestimates MP2 correlation energies.[63,65]

Later, Noga et al.[67] formulated a different approximation for the B intermediate, which is denoted as the standard approximation C. Approximation C is similar to approximation B in that no priori approximations other than RI are made. The only term that is not evaluated directly via an RI insertion is $\langle ij|f_{12}\hat{F}_{12}f_{12}|kl\rangle$, but it can be reformulated with commutators as

$$\begin{aligned}\langle ij|f_{12}\hat{F}_{12}f_{12}|kl\rangle &= \langle ij|f_{12}\big(\hat{F}_{12}+\hat{K}_{12}\big)f_{12}|kl\rangle - \langle ij|f_{12}\hat{K}_{12}f_{12}|kl\rangle \\ &= \frac{1}{2}\langle ij|[f_{12},[\hat{T}_{12},f_{12}]]|kl\rangle + \frac{1}{2}\langle ij|[\hat{F}_{12}+\hat{K}_{12},f_{12}^2]_+|kl\rangle \\ &\quad - \langle ij|f_{12}\hat{K}_{12}f_{12}|kl\rangle,\end{aligned} \tag{3.32}$$

where $[\hat{F}_{12}+\hat{K}_{12},f_{12}^2]_+$ is an anticommutator between $\hat{F}_{12}+\hat{K}_{12}$ and f_{12}^2. In the expression above, the first term can be evaluated analytically, and the second and third terms can be evaluated with the RI approximation. The final expression for the B intermediate in approximation C can be written as

$$\begin{aligned}B_{ij}^{kl} &= \frac{1}{2}\left(\overline{[f_{12}[\hat{T}_{12},f_{12}]]}\right)_{ij}^{kl} \\ &\quad + \frac{1}{2}\left(\left(\overline{f_{12}^2}\right)_{ij}^{p'l}(F+K)_{p'}^{k} + \left(\overline{f_{12}^2}\right)_{ij}^{kp'}(F+K)_{l}^{p'} + \left(\overline{f_{12}^2}\right)_{kl}^{p'j}(F+K)_{p'}^{i} + \left(\overline{f_{12}^2}\right)_{kl}^{ip'}(F+K)_{p'}^{j}\right) \\ &\quad - \bar{r}_{r'p'}^{kl}K_{q'}^{p'}\bar{r}_{ij}^{r'q'} - \bar{r}_{p'm}^{kl}F_{q'}^{p'}\bar{r}_{ij}^{q'm} - \bar{r}_{pa}^{kl}F_{q}^{p}\bar{r}_{ij}^{qa} - \bar{r}_{a'm}^{kl}F_{n}^{m}\bar{r}_{ij}^{a'n} \\ &\quad - \left(\bar{r}_{a'm}^{kl}F_{p'}^{m}\bar{r}_{ij}^{a'p'} + \bar{r}_{pa}^{kl}F_{a'}^{p}\bar{r}_{ij}^{a'a} + \bar{r}_{a'm}^{ij}F_{p'}^{m}\bar{r}_{kl}^{a'p'} + \bar{r}_{pa}^{ij}F_{a'}^{p}\bar{r}_{kl}^{a'a}\right),\end{aligned} \tag{3.33}$$

where p', q', ... are indices for orbitals in the ABS. Similar to V and X, the evaluation of B also involves non-standard two-electron integrals, which include another new type of integrals over $[f_{12}[\hat{T}_{12}, f_{12}]]$. Since there is no need to evaluate the expensive integrals over $[f_{12}, \hat{T}_{12}]$, approximation C requires fewer types of two-electron integrals than those with approximation A and B, and has been shown to perform similarly to approximation B.[2,63] Thus, approximation C represents a simpler way to evaluate the B intermediate than the earlier approaches.

Numerical Quadrature

An alternative approach used to evaluate many-electron integrals is the numerical quadrature. This idea was used by Boys and Handy in their transcorrelated method,[40] and, more recently, employed by Ten-no to compute the many-electron integrals in the MP2-F12 method.[68] The two-electron Coulomb integrals, for example, can be evaluated as sums of two- and three-center objects over grid points:

$$\langle pq|\frac{1}{r_{12}}|rs\rangle = \sum_g w(\mathbf{r}_g)\phi_p(\mathbf{r}_g)\phi_r(\mathbf{r}_g)\langle q|\frac{1}{r_{1g}}|s\rangle, \tag{3.34}$$

where g indexes the grid points, and $w(\mathbf{r}_g)$ is the quadrature weight at the point $\mathbf{r}_g$. Therefore, the transformation of these integrals scales as $\mathcal{O}(N^2OG)$ as opposed to $\mathcal{O}(N^4O)$ in the regular MP2 method, where N is the number of orbitals in the OBS, O is the number of occupied orbitals, and G is the numbers of grid points. If $G \ll N^2$, the scaling is reduced, and thus it is advantageous to use the numerical quadrature. For an accuracy on the order of μE_h, it has been shown that 1000 to 30,000 grid points per atom are required.[68] With the numerical quadrature, other two-electron integrals in the MP2-F12 theory, such as $\langle pq|\hat{T}_1 f(r_{12})|rs\rangle$, can be evaluated in a similar manner.

More importantly, the three-electron integrals in the MP2-F12 theory can be accurately calculated using the numerical quadrature with a reduced scaling. For example, the three-electron integral in Eq. 3.21 can be evaluated directly as

$$\langle ij|\frac{1}{r_{12}}\hat{O}_1 f_{12}|kl\rangle = \sum_{mg} w(\mathbf{r}_g)\phi_j(\mathbf{r}_g)\phi_l(\mathbf{r}_g)\langle i|\frac{1}{r_{1g}}|m\rangle\langle m|f_{1g}|k\rangle. \tag{3.35}$$

Thus, the three-electron integral is decomposed into three-center, one-electron integrals. The computational cost of such evaluation is linear with respect to the number of grid points, whereas the cost is proportional to OP' when the RI approximation is used (see Eq. 3.23), where P' is the number of orbitals in the RI basis. Therefore, the numerical quadrature approach might be more advantageous as the system size increases.[69,70] For more complicated many-electron integrals (e.g., integrals over the exchange operator in the intermediate B), however, a hybrid QD/RI approach is required.[69]

3.1.3 Correlation Factors

In the early days of R12 methods, the linear r_{12} function was used as the correlation factor: $f(r_{12}) = r_{12}$. The explicitly correlated terms that depend linearly on r_{12} give proper description for the wave function at small inter-electronic distances, and thus significantly improve the convergence of correlation energies. However, the linear r_{12} factor yields unphysical long-range correlation behavior: the dynamical correlation between electrons should decay, not grow, at long distances. As a result, the standard expansion of orbital products has to compensate for this unphysical behavior, which means that larger-than-expected basis sets are required in calculations with the linear r_{12} factor.

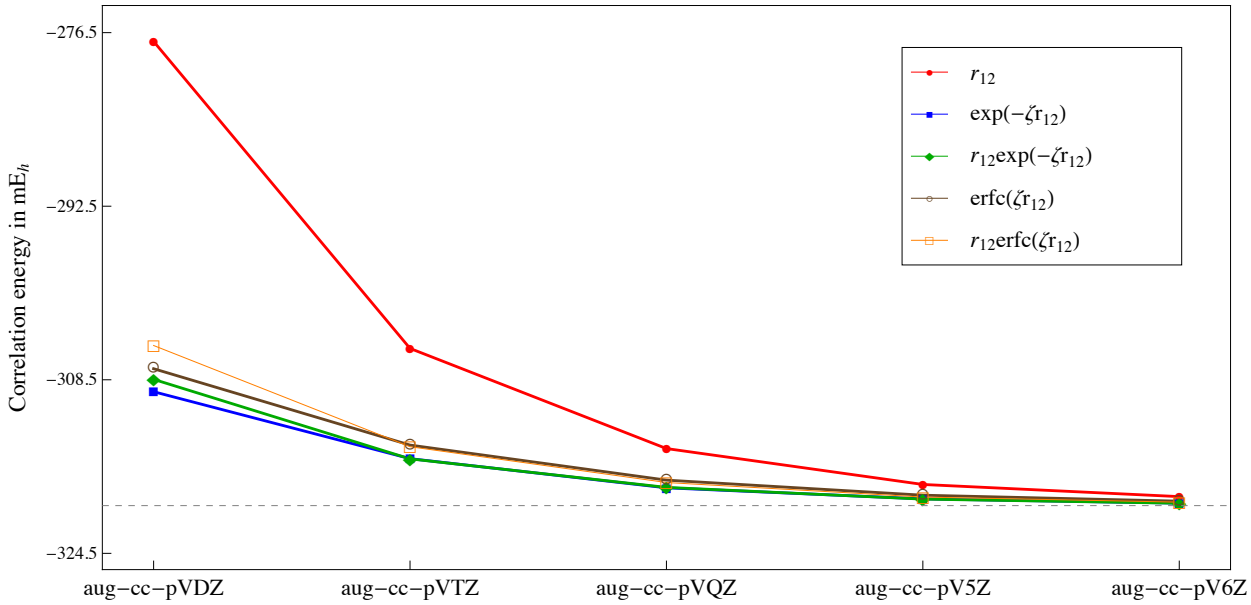

Figure 3.1: Valence MP2 correlation energies (in mE_h) of the Ne atom calculated with MP2-F12(R12) methods using various correlation factors.[3] The optimal ζ varies for different basis sets (see Ref. 3 for more details). The dashed line denotes the basis set limit.

This problem can be solved by using non-linear short-range correlation factors, such as the Slater-type geminal $\exp(-\zeta r_{12})$, which was first proposed by Ten-no,[71] or Gaussian-type geminals, the use of which was investigated by Persson and Taylor in the R12-type framework.[72] The Slater-type geminal gives the correct asymptotes for both $r_{12} \to 0$ and $r_{12} \to \infty$, and thus significantly improves the performance of MP2-R12* calculations with small basis sets.[73] Correlation factors in various other forms, such as $r_{12}\exp(-\zeta r_{12})$, $\mathrm{erfc}(\zeta r_{12})$, and

*Modern applications of R12 methods use the moniker F12 to distinguish from the older, linear r_{12} formulation. It should be noted, however, that there are no other differences between R12 and F12 formalisms. Hence, we will use R12 and F12 interchangeably in this thesis.

Table 3.1: Average unsigned basis set errors (kJ/mol) of HF and various correlation contributions to the reaction energies of the HJO12 set.[a]

Basis	HF	$HF(2)_S$[b]	MP2	MP2-F12	CCSD	$CCSD(2)_{\overline{F12}}$	(T)
aug-cc-pVDZ	13.6	1.2	12.0	2.0	13.1	3.2	1.2
aug-cc-pVTZ	1.1	0.6	3.2	0.8	2.9	0.6	0.4
aug-cc-pVQZ	0.6	0.5	1.3	0.8	1.0	0.6	0.3

[a]The HJO12 set includes 12 isogyric reactions, which range from nearly thermoneutral to highly exothermic and involve breaking and forming an assortment of chemical bonds (see Table 4.2). The complete basis set (CBS) limits for HF energies were the cc-pV6Z HF values, the CBS limits for valence MP2 correlation energies were obtained by subtracting the CBS HF energies and the core-correlation energies from the CBS MP2 energies (obtained by the X^{-3} extrapolation[33] using the cc-pCV5Z and cc-pCV6Z values), and the details of the CBS CCSD and (T) correlation energies can be found in Section 4.3.2. [b]$HF(2)_S$ refers to the Hartree-Fock energies with the CABS singles correction.

$r_{12}\mathrm{erfc}(\zeta r_{12})$,[3] have also been investigated by Tew and Klopper,[3] who compared the performance of MP2-R12 calculations with various correlation factors. Figure 3.1 presents the valence correlation energies of the Ne atom from their calculations, where approximation B was employed, and the correlation-consistent basis sets, aug-cc-pVXZ (X = D, T, Q, 5, 6), were used. They found that the non-linear correlation functions lead to significantly improved results compared to the linear correlation factor: the calculations with them require a basis set of one cardinal number lower than the calculations with r_{12} for an equivalent accuracy. Since the exponent ζ in $\exp(-\zeta r_{12})$ is relatively insensitive to the results, they also concluded that the Slater-type geminal is the best choice.[3] In practice, the Slater-type geminal is usually approximated with a linear combination of Gaussian-type geminals:

$$\exp(-\zeta r_{12}) \approx \sum_{i}^{N} c_i \exp(-\alpha_i r_{12}^2), \tag{3.36}$$

where N is the number of Gaussians. The effect of this approximation on energies is negligible whereas the integrals over the linear combination of Gaussians are easier to evaluate. MP2-F12 methods have been applied to calculate not only energies of closed-shell systems but also energies of open-shell systems[74–76] and molecular properties.[12,77] It has been shown in many benchmark calculations[73,78,79] that MP2-F12 calculations with triple-ζ basis sets yield quintuple-ζ or even better accuracy.

3.1.4 CABS Singles Correction

F12 corrections are very effective in reducing basis set errors of correlation energies, and thus basis set errors of HF energies often become the dominant source of the residual errors

(see Table 3.1 as example). One solution to this problem is to use a large basis set for the computation of the HF energy. Instead of introducing another basis set, the HF basis set error can be greatly reduced perturbatively via standard single excitations into the complementary auxiliary (CA) orbital space. This approach is called "CABS singles," and has been proven to be very effective at reducing HF basis errors.[76,80,81]

In this perturbative approach, the HF determinant is defined as the zeroth-order wave function, and the Hamiltonian can be expressed as:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)} = \begin{pmatrix} F_j^i & 0 & 0 \\ 0 & F_b^a & F_b^{a'} \\ 0 & F_{b'}^a & F_{b'}^{a'} \end{pmatrix} + \begin{pmatrix} 0 & F_i^a & F_i^{a'} \\ F_a^j & 0 & 0 \\ F_{b'}^j & 0 & 0 \end{pmatrix}, \tag{3.37}$$

where a' and b' present orbitals from the CA space. Following the standard perturbation procedure, a second-order energy correction to the HF energy (usually labeled as $E_{(2)_S}$) can be obtained:

$$E_{(2)_S} = E_{HF} + F_i^a t_a^i + F_i^{a'} t_{a'}^i, \tag{3.38}$$

where t_a^i and $t_{a'}^i$ are "CABS singles" amplitudes. This energy correction is not coupled with the MP2 correlation energy. Thus the CABS singles correction can be easily incorporated into MP2-F12 methods. In MP2-F12 calculations, the Fock matrix elements, $F_{b'}^a$ and $F_{b'}^{a'}$, are already available. Therefore, the computational cost for the CABS singles correction is negligible in a MP2-F12 calculation.

In Table 3.1, we list the basis set errors of the HF and various correlation contributions to the electronic reaction energies of the HJO12 set,[1] where the aug-cc-pVXZ (X = D, T, Q) basis sets were used. It is clear that after we include F12 corrections to correlation energies (such as in MP2-F12 and CCSD(2)$_{\overline{F12}}$), the HF basis set errors are the largest source of error especially at the double-ζ level. Nevertheless, the CABS singles correction reduces the HF basis set error to 1.2 kJ/mol at the double-ζ level, and thus it becomes comparable to basis set errors of MP2-F12 correlation energies. The CABS singles correction can also be easily incorporated in the CC-F12 calculations, which ensures that the HF basis set error is not the limiting factor in the calculations. In Table 3.1, we also include the basis set errors of CCSD or (T) correlation energies, and we found that the accuracy of HF(2)$_S$ is also comparable to that of CCSD(2)$_{\overline{F12}}$ or (T) correlation contributions.

3.1.5 Basis sets for F12 Calculations

Despite the tremendous progress of the F12 methods, there are few studies on the performance of various basis sets in F12 calculations. It has been shown that diffuse functions are very important in the F12 calculations,[3] as the electron correlation at the short interelectronic distance can be well described by the geminal functions. Thus, the augmented correlation-consistent basis sets, aug-cc-p(C)VXZ,[18–20] were typically preferred for F12 calculations.

Recently, Peterson and co-workers introduced a new set of correlation-consistent basis sets, cc-p(C)VXZ-F12,[15,82,83] which were optimized specifically for F12 calculations. In their scheme for the cc-pVXZ-F12 basis sets, the contracted HF *s* and *p* functions along with the diffuse *s* and *p* functions were taken from the aug-cc-pV(X+1)Z basis sets. Similarly, the *s*-type correlating functions in the cc-pVXZ-F12 basis sets were simply taken from the cc-pV(X+1)Z basis sets, whereas the *p*-type correlating functions included the two *p* functions from the cc-pV(X+1)Z basis sets and an additional tight *p* function. In the new basis sets, the correlating functions with higher angular momentums along with some *p*-type correlating functions were optimized in the molecular calculations with the MP2-F12 method. The cc-pVXZ-F12 (X = D, T, Q) basis sets were first developed for the atoms H, He, B-Ne, and Al-Ar using a set of closed-shell systems for the basis set optimization. In the process of their optimization, the Slater-type geminal was used as the correlation factor, and thus the optimal geminal exponent was also recommended for each basis set. Later, the optimization was extended to the cc-pVXZ-F12 (X = D, T, Q) basis sets for the alkali and alkaline earth metals: Li, Be, Na, and Mg. Moreover, the core-valence basis sets, cc-pCVXZ-F12 (X = D, T, Q), for the same atoms have also been developed. In these core-valence basis sets, additional functions that account for both core-core and core-valence correlations were included, and they were optimized for the difference between all-electron and valence MP2-F12 correlation energies. Futhermore, they have also reported new optimized auxiliary (OPTRI) basis sets[16,84] for the CABS-based F12 calculations with both the cc-p(C)VXZ-F12 and aug-cc-p(C)VXZ basis sets. These OPTRI basis sets have been shown to have small RI errors and improved numerical stability.

It has been reported that the cc-p(C)VXZ-F12 basis sets give better results than the corresponding aug-cc-p(C)VXZ basis sets in MP2-F12 calculations.[85] However, these calculations are limited to a few properties (e.g., atomization energies). In fact, our work has shown that the performance of these two basis set series vary for CCSD(T)$_{\overline{\text{F12}}}$ calculations on different electronic energy differences (see Chapter 4 for details). Thus, more comprehensive tests of the cc-p(C)VXZ-F12 basis sets are needed.

3.2 CC-R12/F12 Methods

3.2.1 The Formalism

To approach the exact solution of the electronic Schrödinger equation, it is important to improve not only the basis set error with R12 methods but also the accuracy of the underlying method by incorporating high-level correlation effects. The coupled-cluster (CC) theory provides an effective hierarchy of approximations which can, in many cases, rapidly lead to the exact solution. Although the early R12 techniques were developed within the MP2-R12 framework, it is straightforward to apply them to the CC theory. Noga et al. developed CC-R12 methods by extending the standard CC cluster operator ($\hat{T}$) with an additional R12

(geminal) operator ($\hat{R}$).[59,60] The wave function ansatz of CC-R12 methods is defined as

$$|\Psi\rangle = e^{\hat{S}}|\Phi_0\rangle \tag{3.39}$$

with

$$\hat{S} = \hat{T} + \hat{R}. \tag{3.40}$$

The additional R12 operator can be written as

$$\hat{R} = \frac{1}{4} t_{kl}^{ij} \hat{R}_{ij}^{kl} \tag{3.41}$$

with

$$\hat{R}_{ij}^{kl} = \frac{1}{2} \bar{R}_{\alpha\beta}^{kl} \tilde{a}_{ij}^{\alpha\beta}, \tag{3.42}$$

where $\hat{R}_{ij}^{kl}$ can be viewed as a geminal double excitation operator, and $\bar{R}_{\alpha\beta}^{kl}$ are antisymmetrized integrals over the geminal correlation factor. This geminal operator produces two-electron excitations into the complete virtual space, $\{\alpha\}$. The resulting double excitations cannot be generated by the standard $\hat{T}_2$ operator, since $\bar{R}_{\alpha\beta}^{kl}$ are zero when both α and β are in the orbital basis:

$$\bar{R}_{\alpha\beta}^{kl} = \begin{cases} 0 & \text{if } \alpha \in \{\phi_a\} \text{ and } \beta \in \{\phi_b\} \\ \bar{r}_{\alpha\beta}^{kl} & \text{otherwise.} \end{cases} \tag{3.43}$$

With the new wave function ansatz, the Schrödinger equation of CC-R12 methods is written as:

$$\hat{H} e^{\hat{S}} |\Phi_0\rangle = E e^{\hat{S}} |\Phi_0\rangle. \tag{3.44}$$

By multiplying $e^{-\hat{S}}$ along with the reference or excited determinants on its left side, one can obtain the energy and amplitude equations:

$$E = \langle \Phi_0 | e^{-\hat{S}} \hat{H} e^{\hat{S}} | \Phi_0 \rangle, \tag{3.45}$$

$$0 = \langle \Phi_0 | \big(\tilde{a}_{ij\ldots}^{ab\ldots}\big)^\dagger e^{-\hat{S}} \hat{H} e^{\hat{S}} | \Phi_0 \rangle, \tag{3.46}$$

$$0 = \langle \Phi_0 | \big(\hat{R}_{ij}^{kl}\big)^\dagger e^{-\hat{S}} \hat{H} e^{\hat{S}} | \Phi_0 \rangle. \tag{3.47}$$

In the coupled-cluster singles and doubles R12 (CCSD-R12) model, the energy equation can be expressed as

$$\begin{aligned} E_{\text{CCSD-R12}} &= \langle \Phi_0 | \bar{H}_{\text{CCSD}} + [\bar{H}_{\text{CCSD}}, \hat{R}] | \Phi_0 \rangle \\ &= \langle \Phi_0 | \bar{H}_{\text{CCSD}} | \Phi_0 \rangle + \langle \Phi_0 | [\hat{W}, \hat{R}] | \Phi_0 \rangle \\ &= E_{\text{CCSD}} + \frac{1}{4} t_{kl}^{ij} \bar{V}_{ij}^{ab}, \end{aligned} \tag{3.48}$$

where $\bar{H}_{\text{CCSD}}$ is the CCSD similarity-transformed Hamiltonian ($\bar{H}_{\text{CCSD}} = e^{-\hat{T}_1 - \hat{T}_2} \hat{H} e^{\hat{T}_1 + \hat{T}_2}$), $\hat{F}$ is the Fock operator, and $\hat{W}$ is the fluctuation operator defined as $\hat{W} \equiv \hat{H} - \hat{F}$. With

optimized CCSD and geminal amplitudes, the CCSD-R12 energy can be easily computed. These amplitudes can be obtained by solving the following equations:

$$0 = \langle \Phi_i^a | \bar{H}_{\mathrm{CCSD}} + [\bar{H}_{\mathrm{CCSD}}, \hat{R}] | \Phi_0 \rangle, \tag{3.49}$$

$$0 = \langle \Phi_{ij}^{ab} | \bar{H}_{\mathrm{CCSD}} + [\bar{H}_{\mathrm{CCSD}}, \hat{R}] + \frac{1}{2}[[\bar{H}_{\mathrm{CCSD}}, \hat{R}], \hat{R}] | \Phi_0 \rangle, \tag{3.50}$$

$$0 = \langle \Phi_{ij}^{kl} | \bar{H}_{\mathrm{CCSD}} + [\bar{H}_{\mathrm{CCSD}}, \hat{R}] + \frac{1}{2}[[\bar{H}_{\mathrm{CCSD}}, \hat{R}], \hat{R}] | \Phi_0 \rangle, \tag{3.51}$$

where $\langle \Phi_{ij}^{kl}| \equiv \langle \Phi_0 | \big(\hat{R}_{ij}^{kl}\big)^\dagger$. However, the geminal amplitude equation (Eq. 3.51) introduces an additional computational step, which scales as $\mathcal{O}(O^3V'^3)$ (O is the number of occupied orbitals and V' is the number of CABS orbitals). Moreover, two new F12 intermediates,

$$P_{kl}^{ij} = \bar{R}_{kl}^{\alpha\beta} \bar{g}_{\alpha\beta}^{\gamma\delta} \bar{R}_{\gamma\delta}^{ij}, \tag{3.52}$$

$$Z_{kl;m}^{ij;p} = \bar{R}_{kl}^{\alpha\beta} \bar{g}_{\alpha m}^{\gamma p} \bar{R}_{\gamma\beta}^{ij}, \tag{3.53}$$

are required in the evaluation,[59] where $ij;p$ and $kl;m$ indicate that $Z_{kl;m}^{ij;p}$ is antisymmetric with respect to the permutation of i and j as well as k and l. Unlike all other terms, these two intermediates have to be evaluated analytically since direct insertions of the RI result in integrals that converge slowly with respect to the size of the CABS. Thus, their evaluation leads to a significant increase in the computational cost.

Compared to the standard CC methods, the implementation of the CC-R12 methods is much more complicated due to the R12 terms. Thus, the early implementation of CC-R12 methods utilized the SA, which significantly simplified the equations and software implementations. A large basis set, however, is required to approximate the RI. Therefore, they are mostly used for highly accurate computations of atoms and small molecules (up to five atoms).[86,87] Full implementations of CC-R12 methods with the CABS approach have become possible with automated equation derivation and implementation tools. Shiozaki et al.[88,89] reported the first implementation of the CCSD-R12 method using the symbolic algebra code SMITH, and later they extended the implementation to higher-rank CC-R12 methods,[90] such as the CCSDT-R12 and CCSDTQ-R12 methods. A similar implementation of the full CCSD-R12 method has also been reported by Köhn et al.[91] with the GECCO program.

3.2.2 The SP Approach

Instead of solving geminal amplitude equations, an alternative is to use fixed geminal amplitudes as proposed by Ten-no (called the SP approach).[68,71,92] In this approach, the geminal amplitudes are determined by the s-wave (singlet) and p-wave (triplet) cusp conditions, as these conditions describe the behavior of the exact wave function at small inter-electronic distances (see Eq. 1.33 and 1.35). Therefore, the $\hat{R}$ operator in the SP approach can be written as[93]

$$\hat{R}(\mathrm{SP}) = \frac{1}{2} \bar{\mathcal{R}}_{\alpha\beta}^{ij} \tilde{a}_{ij}^{\alpha\beta} \tag{3.54}$$

with

$$\bar{\mathcal{R}}_{\alpha\beta}^{ij} = \langle \alpha\beta | \hat{Q}_{12} \big(\frac{1}{2}\hat{P}_0 + \frac{1}{4}\hat{P}_1 \big) f_{12} | ij \rangle, \tag{3.55}$$

Here, $\hat{P}_0$ and $\hat{P}_1$ are the singlet and triplet spin projectors:

$$\hat{P}_0 = |\alpha\beta\rangle_0 \langle \alpha\beta|_0, \tag{3.56}$$

$$\hat{P}_1 = |\beta\beta\rangle\langle\beta\beta| + |\alpha\beta\rangle_1 \langle \alpha\beta|_1 + |\alpha\alpha\rangle\langle\alpha\alpha|, \tag{3.57}$$

where

$$|\alpha\beta\rangle_0 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\}, \tag{3.58}$$

$$|\alpha\beta\rangle_1 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) + \beta(1)\alpha(2)\}. \tag{3.59}$$

These projectors are used to ensure the appropriate selection of the coefficient for the singlet and triplet pairs.[†] In addition, the contributions from F12 terms in the SP approach can be pre-evaluated due to the fixed amplitudes, and thus further reduce its computational cost.

The use of fixed geminal amplitudes introduces errors into the energy. Therefore, the Lagrange functional is usually used to compute the energy in the SP approach, as the error in the energy depends quadratically on the error in the geminal amplitudes. For example, the CCSD-F12 Lagrangian can be written as

$$L_{\text{CCSD-F12}} = L_{\text{CCSD}} + L_{\text{F12}}, \tag{3.60}$$

$$L_{\text{CCSD}} = E_0 + \langle \Phi_0 | \bar{H}_{\text{CCSD}} | \Phi_0 \rangle + \langle \Phi_0 | \hat{\Lambda}_1 \bar{H}_{\text{CCSD}} | \Phi_0 \rangle + \langle \Phi_0 | \hat{\Lambda}_2 \bar{H}_{\text{CCSD}} | \Phi_0 \rangle, \tag{3.61}$$

$$\begin{aligned} L_{\text{F12}} = {} & \langle \Phi_0 | [\hat{H}_{\text{CCSD}}, \hat{R}] | \Phi_0 \rangle + \langle \Phi_0 | \hat{\Lambda}_1 [\bar{H}_{\text{CCSD}}, \hat{R}] | \Phi_0 \rangle \\ & + \langle \Phi_0 | \hat{\Lambda}_2 \big([\bar{H}, \hat{R}] + \frac{1}{2}[[\bar{H}_{\text{CCSD}}, \hat{R}], \hat{R}] \big) | \Phi_0 \rangle \\ & + \langle \Phi_0 | \hat{\Lambda}_2' \big([\bar{H}_{\text{CCSD}}, \hat{R}] + \frac{1}{2}[[\bar{H}_{\text{CCSD}}, \hat{R}], \hat{R}] \big) | \Phi_0 \rangle, \end{aligned} \tag{3.62}$$

where $\hat{\Lambda}$ operators are defined as

$$\hat{\Lambda}_1 = \lambda_i^a \tilde{a}_a^i, \tag{3.63}$$

$$\hat{\Lambda}_2 = \frac{1}{(2!)^2} \lambda_{ij}^{ab} \tilde{a}_{ab}^{ij}, \tag{3.64}$$

$$\hat{\Lambda}_2' = \frac{1}{(2!)^2} \lambda_{ij}^{kl} (\hat{R}_{ij}^{kl})^\dagger. \tag{3.65}$$

[†]It should be noted that this requires f_{12} to behave linearly at short r_{12}, which means that for the exponential factor we must set $f_{12} = (1 - \exp(-\zeta r_{12}))/-\zeta$, so that its Taylor expansion at small r_{12} is $f_{12} \approx r_{12} + \mathcal{O}(r_{12}^2)$.

By making the Lagrangian stationary with respect to the Lagrangian multipliers (λ), the optimized amplitudes can be obtained. In the SP approach, $\hat{R}$ is set to $\hat{R}(\mathrm{SP})$, and Λ_2' is equal to $[\hat{R}(\mathrm{SP})]^\dagger$.[4] As expected, the reduced flexibility due to the fixed geminal amplitudes results in a loss of accuracy in total energies computed with the SP approach. The accuracy of relative energies, however, is often comparable to that of CC-R12 methods with optimized geminal amplitudes. Moreover, the SP approach is numerically more stable, and is also free of the geminal basis set superposition error, which leads to improved predictions for weak interactions.[92] As a result, the SP approach has been adopted in most practical applications.[2,10]

3.2.3 Approximate CC-R12/F12 Methods

To extend applications to larger systems, various simplified CC-R12/F12 models, which demand less computational effort while maintaining similar accuracy as the CC-R12/F12 methods, have been developed over the years. These methods include CC(R12/F12),[94] CC(F12*),[95] CC-F12x (x = a, b),[80,96] and CC(2)$_{\overline{\mathrm{F12}}}$.[97,98] Their performance has been closely examined in the work of Hättig et al.[95] In Table 3.2, we present parts of their data that include the mean and maximum basis set errors of reaction and atomization energies calculated with various approximate CCSD-F12 models. While the performance of these approximate CCSD-F12 methods is similar in general, there are some small variations for calculations of different properties. In the following, we will present more details of these models, and further discuss their differences.

CC(F12) Methods

The CCSD(R12) method was first introduced by Klopper and co-workers,[94] and they first implemented the method in the DALTON program.[99] Today, the implementation of the CC(R12/F12) methods can also be found in many other scientific software packages like TURBOMOLE[100] and MPQC[101] packages. In the CCSD(R12) method, only the terms linear in geminal amplitudes are retained in the amplitude equations (Eq. 3.50 and Eq. 3.51). Assuming the orthogonality between geminal and conventional double excitations is satisfied, the nonlinear terms are very small. Furthermore, the $[\bar{H}, \hat{R}]$ term is reduced to $[\hat{F}, \hat{R}]$ in the geminal amplitude equation (Eq. 3.51), which eliminates the calculation of small terms that are computationally expensive and numerically inaccurate. As a result, the double excitation and geminal amplitude equations is transformed to:

$$0 = \langle \Phi_{ij}^{ab} | \bar{H}_{\mathrm{CCSD}} + [\bar{H}_{\mathrm{CCSD}}, \hat{R}] | \Phi_0 \rangle, \tag{3.66}$$

$$0 = \langle \Phi_{ij}^{kl} | \bar{H}_{\mathrm{CCSD}} + [\hat{F}, \hat{R}] | \Phi_0 \rangle \tag{3.67}$$

in the CCSD(R12) method, and the corresponding correlation energy is computed with

$$E_{\mathrm{CCSD(R12)}} = \langle \Phi_0 | \bar{H}_{\mathrm{CCSD}} + [\bar{H}_{\mathrm{CCSD}}, \hat{R}] | \Phi_0 \rangle. \tag{3.68}$$

Table 3.2: The mean and maximum basis set errors (kJ/mol per valence electron) of various CCSD correlation contributions to reaction energies and atomization energies.[4,a]

Method	Basis	Reaction energy		Atomization energy	
		Mean	Max	Mean	Max
CCSD(F12)	cc-pVDZ-F12	−0.10	−0.32	−0.61	−1.05
	cc-pVTZ-F12	−0.01	−0.06	−0.08	−0.33
	cc-pVQZ-F12	0.00	0.01	0.04	−0.13
CCSD(F12*)	cc-pVDZ-F12	−0.10	−0.32	−0.60	−1.04
	cc-pVTZ-F12	−0.01	−0.06	−0.08	−0.33
	cc-pVQZ-F12	0.00	0.01	0.04	−0.13
CCSD[F12]	cc-pVDZ-F12	−0.09	−0.30	−0.32	−0.76
	cc-pVTZ-F12	−0.01	−0.05	0.05	−0.26
	cc-pVQZ-F12	0.00	−0.01	0.10	0.17
CCSD-F12a	cc-pVDZ-F12	−0.14	−0.50	−0.37	−0.70
	cc-pVTZ-F12	−0.02	−0.11	0.19	0.58
	cc-pVQZ-F12	0.00	0.02	0.27	0.45
CCSD-F12b	cc-pVDZ-F12	−0.16	−0.53	−0.85	−1.54
	cc-pVTZ-F12	−0.03	−0.12	−0.16	−0.40
	cc-pVQZ-F12	0.00	−0.03	0.05	0.14
CCSD(2)$_{\overline{\text{F12}}}$	cc-pVDZ-F12	−0.13	−0.42	−0.72	−1.48
	cc-pVTZ-F12	−0.02	−0.08	−0.13	−0.53
	cc-pVQZ-F12	0.00	0.01	0.04	−0.19

[a]The test set for atomization energies include 30 small closed-shell molecules that contains H, C, N, O, and F elements, and the decomposition of these 25 molecules into H_2, CO, CO_2, N_2, and F_2 comprises the reaction test set. In the calculations, the SP approach is used.

If the SP approach is used, however, the energy is evaluated using the Lagrangian instead of Eq. 3.68.

Later, the CCSD(R12) approach has also been extended to include F12 geminal functions as the correlation factor,[102,103] which decays exponentially with the inter-electronic distance, and such method is referred to as the CCSD(F12) method. Within the SP approximation, the most time consuming process during a CCSD(F12) calculation is the iterative computation of the CCSD amplitudes. In these iterations, most terms contributed from F12 geminals are less expensive than the conventional terms. On the other hand, CCSD(F12) calculations produce correlation energies of quintuple-ζ quality with triple-ζ basis sets.[102,103] Thus, the CCSD(F12) method represents an excellent approximation to the CCSD-F12 method: correlation energies close to the CCSD basis set limit can be obtained with a relatively low computational cost. The CCSD(T)(R12)[104] and CCSD(T)(F12),[102,103] methods have also been developed, where the triples corrections are computed in the same way as those in the standard CCSD(T) theory. With these methods, chemical accuracy in energy differences can be obtained at a reasonable cost.

Recently, simplified CCSD(F12) methods, CCSD(F12*) and CCSD[F12], have also been reported by Hättig and co-workers.[95] Within the SP approach, the additional computational cost in these methods is further reduced, and is now comparable to that of CCSD-F12b and CCSD(2)$_{\overline{\text{F12}}}$ methods (see the following sections for details). On the other hand, it has been shown that the simplifications introduced in these two methods result in negligible errors in energies.[4] In Figure 3.2, for example, the CCSD(F12*) and CCSD[F12] methods show equivalent accuracy to the CCSD(F12) method for the calculations of reaction energies and atomization energies within the SP approach.

CC-F12x Methods

The CC-F12x (x = a, b) methods were developed by Werner and co-workers,[80,96] who implemented these methods in the MOLPRO package.[105] In the CC-F12x methods, the SP ansatz is always used. As a result, it is no longer necessary to solve the geminal amplitude equations, and thus the number of equations is the same as that in standard CC methods. Furthermore, the double excitation amplitude equation is reduced to

$$0 = \langle \Phi_{ij}^{ab} | \bar{H}_{\text{CCSD}} + [\hat{H}, \hat{R}] | \Phi_0 \rangle, \tag{3.69}$$

where the $[[\hat{W}, \hat{T}_2], \hat{R}]$ term included in Eq. 3.66 is neglected. In addition, the contributions from the CABS orbitals are also neglected except those present in the MP2-F12 theory. For the two CC-F12x variants, the difference lies in their energy expressions:

$$E_{\text{CCSD-F12a}} = E_{\text{CCSD}} + \langle \Phi_0 | \hat{\Lambda}_2' \big(\hat{W} + [\hat{F}, \hat{T}_2] + [\hat{F}, \hat{R}] \big) | \Phi_0 \rangle, \tag{3.70}$$

$$E_{\text{CCSD-F12b}} = E_{\text{CCSD}} + \langle \Phi_0 | \hat{\Lambda}_2' \big(\hat{W} + [\hat{H}, \hat{T}_2] + [\hat{F}, \hat{R}] \big) | \Phi_0 \rangle, \tag{3.71}$$

where an additional coupling term, $[\hat{W}, \hat{T}_2]$, is included in the CCSD-F12b energy expression. Due to the evaluation of the additional F12 terms in the CC amplitude equations, each CC iteration in the CCSD-F12x methods takes a small amount of extra time. However, the predominant additional cost in CCSD-F12x calculations is an initial MP2-F12 calculation, which scales as $\mathcal{O}(N^5)$ (N is a measure of the molecular size). Compared to the CCSD method, which scales as $\mathcal{O}(N^6)$, this additional computational cost is relatively small. In CCSD(T)-F12x calculations, it becomes negligible for larger molecules, as the scaling of the CCSD(T) method is $\mathcal{O}(N^7)$.[10]

CC-F12x methods can be considered approximations to the CC(F12) approach, which retain only the dominant F12 contributions in the amplitude equations. While they are slightly less accurate than the CC(F12) models, their computational cost is lower. It has been shown that the CCSD-F12a method slightly overestimates the CCSD-F12 correlation energies whereas the CCSD-F12b variant slightly underestimates them. However, relative energies computed with these two methods have similar accuracy.[106] For example, the CCSD-F12a and CCSD-F12b models show comparable accuracy for calculations on reaction energies (see Table 3.2). For calculations on atomization energies, the CCSD-F12a model gives better results than the CCSD-F12b variant with the cc-pVDZ-F12 basis set. With larger basis sets, however, its yields slightly worse results than CCSD-F12b.

CC(2)$_{\overline{\mathbf{F12}}}$ Methods

The CC(2)$_{\overline{\mathrm{F12}}}$ models were first proposed by Valeev and implemented within the MPQC package,[101] and now can also be found in other packages, such as MOLPRO.[105] In the CC(2)$_{\overline{\mathrm{F12}}}$ methods,[97,98] the geminal contributions are included with the perturbation theory. This approach is analogous to the way in which the connected triples in the CCSD(T) method are treated.[107] In the CCSD(2)$_{\overline{\mathrm{F12}}}$ model, the Lagrangian is rewritten in a matrix form as

$$L_{\text{CCSD-F12}} = \mathbf{L}^{\dagger}\bar{\mathbf{H}}\mathbf{R}, \tag{3.72}$$

where $\mathbf{L}$ and $\mathbf{R}$ are the left- and right-hand eigenvectors of $\bar{\mathbf{H}}$, and $\bar{\mathbf{H}}$ is the matrix representation of the similarity-transformed CCSD Hamiltonian ($\bar{H} \equiv e^{(-\hat{T}_1-\hat{T}_2)}\hat{H}e^{(\hat{T}_1+\hat{T}_2)}$). This Hamiltonian matrix can be written as

$$\bar{\mathbf{H}} = \begin{pmatrix} \bar{\mathbf{H}}_{\mathrm{PP}} & \bar{\mathbf{H}}_{\mathrm{PQ}} \\ \bar{\mathbf{H}}_{\mathrm{QP}} & \bar{\mathbf{H}}_{\mathrm{QQ}} \end{pmatrix}, \tag{3.73}$$

where P is the reference space including the reference determinant (0), singly substituted determinants (S), and doubly substituted determinants (D); and Q represents an external space comprised of explicitly correlated geminal substitutions (Γ). With a Löwdin-type perturbation expansion,[108] the matrices of the zeroth- and first-order Hamiltonians can be

obtained by partitioning the similarity-transformed Hamiltonian matrix:

$$\bar{\mathbf{H}}^{(0)} = \begin{pmatrix} \bar{\mathbf{H}}_{\mathrm{PP}} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{H}}^{(0)}_{\mathrm{PP}} \end{pmatrix}, \tag{3.74}$$

$$\bar{\mathbf{H}}^{(1)} = \begin{pmatrix} \mathbf{0} & \bar{\mathbf{H}}_{\mathrm{PQ}} \\ \bar{\mathbf{H}}_{\mathrm{QP}} & \bar{\mathbf{H}}^{(1)}_{\mathrm{QQ}} \end{pmatrix}, \tag{3.75}$$

where the zeroth-order Hamiltonian is the standard Fock operator ($\hat{F}$), and the first-order Hamiltonian is the similarity-transformed CCSD Hamiltonian ($\bar{H}$). Thus, the left- and right-hand zeroth-order eigenvectors in the CCSD(2)$_{\overline{\mathrm{F12}}}$ method are the standard left- and right-hand ground-state CCSD eigenvectors ($\mathcal{L}$ and $\mathcal{R}$),[1]

$$\mathbf{L}^{(0)} = \begin{pmatrix} \mathcal{L} \\ \mathbf{0} \end{pmatrix}, \tag{3.76}$$

$$\mathbf{R}^{(0)} = \begin{pmatrix} \mathcal{R} \\ \mathbf{0} \end{pmatrix}. \tag{3.77}$$

and the zeroth-order energy is the ground-state CCSD energy. The geminal functions constitute the first-order wave function, and thus lead to a second-order energy correction.

In the CCSD(2)$_{\overline{\mathrm{F12}}}$ method, CC amplitude equations are not modified. As a result, the computational cost of the iterations for CCSD amplitudes in this method is identical to that in the standard CCSD method. The additional computational cost comes from the evaluation of F12 intermediates and the final energy. Valeev employed the screening approximations (ScrA) in the CCSD(2)$_{\overline{\mathrm{F12}}}$ model to simplify the formula,[97] and those approximations lead to a second-order Hylleraas-type functional from F12 contributions:

$$L_{(2)_{\overline{\mathrm{F12}}}} = \frac{1}{8} t_{ij}^{ow} (\tilde{B}_{(ij)})_{ow}^{kl} t_{kl}^{ij} + \frac{1}{2} \bar{V}_{ij}^{kl} t_{kl}^{ij}. \tag{3.78}$$

Compared with the MP2-F12 Lagrangian, the difference in $L_{(2)_{\overline{\mathrm{F12}}}}$ lies in the V intermediate. The tensor element of the V intermediate here can be expressed as

$$\bar{V}_{ij}^{kl} \equiv V_{ij}^{kl} + \frac{1}{2} \left(V_{ab}^{kl} + C_{ab}^{kl} \right) t_{ij}^{ab} + V_{aj}^{kl} t_i^a + V_{ia}^{kl} t_j^a, \tag{3.79}$$

where not only the first term, which contributes to the V intermediate in the MP2-F12 theory, is included, but also terms that involve the CCSD amplitudes, t_i^a and t_{ij}^{ab}.

The error introduced by the approximations in the CCSD(2)$_{\overline{\mathrm{F12}}}$ model are very small, and are negligible when compared to the residual basis set error of the CCSD-F12 method. It has been shown that the CCSD(2)$_{\overline{\mathrm{F12}}}$ method is a very good approximation to the CCSD-F12 method.[2,10] The CCSD(2)$_{\overline{\mathrm{F12}}}$ method is slightly worse than the CCSD(F12) models, but generally more accurate than the CCSD-F12x methods (as shown in Figure 3.2). In addition, Valeev et al. have also extended the CC(2)$_{\overline{\mathrm{F12}}}$ approach to the CCSD(T) level, which includes

the effect from connected three-electron correlations.[109] In their CCSD(T)$_{\overline{\text{F12}}}$ method, the geminal and triples effects are completely decoupled. Moreover, the SP approximation is typically employed in their calculations.[93,98] Thus, the CCSD(T)$_{\overline{\text{F12}}}$ energy can be simply computed as

$$E_{\text{CCSD(T)}_{\overline{\text{F12}}}} = E_{\text{CCSD}} + E_{(\text{T})} + E_{(2)_{\text{S}}} + E_{(2)_{\overline{\text{F12}}}}, \tag{3.80}$$

which is the sum of the standard CCSD(T) energy, the usual "CABS singles" correction, and the F12 correction.

3.3 Recent Developments

3.3.1 Multi-Reference F12 Methods

Until recently, most of the F12 methods were based on the assumption that a single reference is sufficient to approximate the wave function. However, there are systems where the multi-reference (MR) treatment is necessary for the proper description of the system. For example, the calculations of excited states and bond breakings require the multi-reference methods. Gdanitz[61,62] was the first to develop the MR configuration interaction R12 method, where the conventional MRCI wave function is augmented with explicitly correlated terms:

$$\Psi = \Psi_{\text{MRCI}} + \sum_I \hat{Q}^I t_{kl}^{ij}(I) r_{\kappa\lambda}^{kl} a_{ij}^{\kappa\lambda} |\Psi_I\rangle \tag{3.81}$$

where Ψ_I represent the reference determinants, and the $\hat{Q}^I$ projector, which depends on the Ith reference determinant, ensures the orthogonality between the explicitly correlated functions and the conventional MRCI wave function. It should be noted that the geminal amplitudes also depend on the reference determinants, which results in a large number of parameters to be optimized. He also developed the explicitly correlated MR averaged coupled-pair functional, which has been used to study various properties (e.g., potential energy surfaces and excited states) of small systems.

More recently, Ten-no[75] proposed an R12 version of the MR-MP2 method, where the explicitly correlated terms are included in an internally contracted manner. In his approach, the first-order wave function is written as

$$\Psi^{(1)} = \Psi_{\text{MRMP}}^{(1)} + \hat{Q}\hat{R}_{12} \sum_I t_I |\Psi_I\rangle \tag{3.82}$$

with

$$\hat{R}_{12} = \frac{1}{2} t_{kl}^{ij} r_{\kappa\lambda}^{kl} a_{ij}^{\kappa\lambda}, \tag{3.83}$$

where explicitly correlated functions are generated by $\hat{R}_{12}$ when acting on the reference determinants. Ten-no's approach not only leads to a more compact formula, but also reduces

the number of geminal amplitudes to be determined as they do not directly depend on reference determinants. In Ten-no's work, the SP approach was also adopted, which further reduced the computational cost of the method and improved its numerical stability. In addition, Torheyden and Valeev[110] have presented a general second-order R12 correction, $[2]_{\mathrm{R12}}$, which can be applied to any reference state as long as the two-electron reduced density matrix is available. The F12 versions of the CASPT2 and MRCI methods have also been presented by Shiozaki and Werner.[111,112] For more details of these methods, we refer interested readers to the review paper from Valeev et al.[2]

3.3.2 Monte Carlo F12 Method

Recently, Hirata et al.[113] have implemented a Monte Carlo explicitly correlated MP2 (MC-MP2-F12) method, where a Monte Carlo (MC) integration was used to compute the sum of many-electron integrals. In general, the integral I of a high-dimensional function, $f(\mathbf{x})$, can be approximated as the sum of the quotient $f(\mathbf{x}_n)/w(\mathbf{x}_n)$,

$$\begin{aligned} I &= \iiint d\mathbf{x} f(\mathbf{x}) \\ &\approx \frac{1}{N}\sum_{n=1}^{N}\frac{f(\mathbf{x}_n)}{w(\mathbf{x}_n)} \end{aligned} \tag{3.84}$$

where $\mathbf{x}_n$ are random sampling points distributed according to the weight function, $w(\mathbf{x})$, and N is the number of sampling points.[114] In a similar manner, the F12 energy corrections of the MP2-F12 method with Ten-no's fixed amplitudes[71] (see Section 3.2.2 for more details),

$$E_{\mathrm{F12}} = \frac{5}{8}\sum_{i,j}^{\mathrm{occ.}}\langle ij|r_{12}^{-1}\hat{Q}_{12}f_{12}|ij\rangle - \frac{1}{8}\sum_{i,j}^{\mathrm{occ.}}\langle ij|r_{12}^{-1}\hat{Q}_{12}f_{12}|ji\rangle, \tag{3.85}$$

can be evaluated with the MC integration. For example, the F12 energy contribution from the three-electron integrals can be written as

$$E_{\mathrm{3e}} = -\frac{5}{4}\sum_{i,j,k}^{\mathrm{occ.}}\langle ijk|\frac{f_{23}}{r_{12}}|kji\rangle + \frac{1}{4}\sum_{i,j,k}^{\mathrm{occ.}}\langle ijk|\frac{f_{23}}{r_{12}}|kij\rangle. \tag{3.86}$$

With the MC integration, it can be evaluated as

$$\begin{aligned} E_{\mathrm{3e}} &= \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3\, F_3(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \\ &\approx \frac{1}{N}\sum_{n=1}^{N}\frac{F_3(\mathbf{r}_1^{[n]},\mathbf{r}_2^{[n]},\mathbf{r}_3^{[n]})}{w_3(\mathbf{r}_1^{[n]},\mathbf{r}_2^{[n]},\mathbf{r}_3^{[n]})}, \end{aligned} \tag{3.87}$$

Table 3.3: The F12 corrections (E_{F12} in mE_{h}) from RI- and MC-MP2-F12 calculations and the statistical uncertainty (ε in mE_{h}) in MC-MP2-F12 calculations.[a]

Molecule	Basis	E_{F12}(RI)	E_{F12}(MC)	ε
H_2O	cc-pVDZ	−82.2	−82.1	0.3
H_2O	cc-pVTZ	−35.0	−34.5	0.7
H_2O	cc-pVQZ	−16.1	−16.3	1.6
CH_4	cc-pVDZ	−46.0	−46.2	0.3
CH_4	cc-pVTZ	−18.5	−19.3	1.1
CH_4	cc-pVQZ	−8.3	−9.2	2.6
C_6H_6	cc-pVDZ	−215.4	−214.7	1.3
C_6H_6	cc-pVTZ	−95.0	−94.9	4.3
C_6H_6	cc-pVQZ	−44.5	−35.8	9.9

[a]The Hylleraas functional was used in the RI-MP2-F12 calculations. In all calculations, the geminal exponent, $\gamma = 1.2$ was employed. For H_2O and CH_4, 2×10^6 MC steps were used, while 8×10^6 steps were used for C_6H_6.

where the integrand is defined by

$$F_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = -\frac{5}{4}\frac{f_{23}O_{13}O_{22}O_{31}}{r_{12}} + \frac{1}{4}\frac{f_{23}O_{12}O_{23}O_{31}}{r_{12}} \tag{3.88}$$

with

$$O_{pq} = \sum_{i}^{\text{occ.}} \varphi_i^*(\mathbf{r}_p)\varphi_i(\mathbf{r}_q). \tag{3.89}$$

Table 3.3 lists the F12 energy corrections from the MC-MP2-F12 calculations and the MP2-F12 calculations for a few molecules with the cc-pVXZ (X = D, T, Q) basis sets.[18] In our MP2-F12 calculations, the RI approximation was used in the evaluation of many-electron integrals. It is clear that the MC-MP2-F12 method can give results that are within a few mE_{h} of the MP2-F12 method. However, the uncertainty of the MC-MP2-F12 calculation increases as the size of the basis set increases. Nevertheless, the MC-MP2-F12 method presents a viable alternative to the regular MP2-F12 method when small basis sets are used. Since the operation cost of the MC integration per step scales quadratically with size, the MC-MP2-F12 method might be advantageous for large-scale systems.

In addition, Flad et al.[115] have studied the performance of the transcorrelated equation for the quantum Monte Carlo (QMC) Jastrow factors, the product of which with the reference comprises the wave function. In their calculations, the integrals involved are evaluated using QMC methods. Their test calculations on a few atoms and molecules demonstrated that QMC can be an efficient approach to compute the complicated integrals involved in explicitly correlated methods. More recently, they proposed a general approach to combine

QMC and F12 methods,[116] where optimized pair-correlation functions from Jastrow factors are transferred into F12 methods.

3.4 Summary

The F12 approach reduces the basis set error of the correlation energy by augmenting the orbital expansion with a small set of explicitly correlated geminals, which can efficiently describe the wave function near the electron coalescence. The F12 methods were first implemented by Kutzelnigg and Klopper within the MP2 framework.[58,64,117] Since then, the rapid improvement of the F12 technology has significantly improved the efficiency and accuracy of F12 methods. The key developments include robust RI approximations,[66,67] numerical quadrature,[68] density fitting,[118,119] and efficient nonlinear correlation factors.[71,73,79] New basis sets specifically designed for F12 methods[15–17] have also been developed. However, their performance with F12 calculations needs to be investigated with more comprehensive tests. F12 corrections have been shown to be very effective in reducing basis set errors of correlation energies, and the HF basis set errors are often the largest source of the residual errors. Thus, the CABS singles correction is also included in the F12 calculations to reduce the HF basis errors. The basis set errors of HF energies with the CABS singles correction are comparable to those of correlation energies with F12 corrections.

To further improve the accuracy of the F12 calculations, coupled-cluster F12 methods[59,88,91] were introduced. Due to the complexity of the full CC-R12 methods, practical approximations to CC-F12,[76,80,94,95,97,109] both iterative and non-iterative, have also been developed. These approximate CC-F12 methods are almost as accurate as the full CC-F12 methods, but their computational cost is a fraction of the cost of the CC-F12 methods. Among them, the CC(F12) methods are the most expensive and usually the most accurate approximate CC-F12 methods. The CC(F12*) and CC[F12] variants are simplified CC(F12) methods, but retain equivalent accuracy to the CC(F12) models. While less expensive, CC-F12x and CC(2)F12 models are slightly less accurate than the CC(F12) methods. In general, their performance is similar with some small variations for calculations of different properties. Today the F12 methods have become efficient and competitive with the extrapolation approach:[120] approximate CC-F12 methods require a basis set of two cardinal numbers lower compared with the standard CC methods for the same accuracy.

Despite these tremendous developments, the F12 approach continues to evolve. Recently, it has been extended to study systems where multi-reference treatment is needed.[62,75,110,111] Moreover, new F12 variants, such as MC-MP2-F12,[113] have begun to emerge. In addition, recent work has been extended to computations of response properties[13,14,121] and relativistic effects.[122,123] In the future, we can reasonably expect more developments in these areas.

Chapter 4

Prediction of Reaction Barriers and Thermochemical Properties with Explicitly Correlated Coupled-Cluster Methods: A Basis Set Assessment

4.1 Introduction

Quantitative low-temperature computational kinetics are challenging because they require predictions of electronic energy differences (reaction energies, atomization energies, enthalpies of formation, and reaction barriers) with chemical accuracy, usually defined as errors not exceeding 1 kcal/mol (4.18 kJ/mol) or even as 1 kJ/mol. Despite the impressive recent progress of density functional theory (DFT) methodology, its applicability to computational kinetics is still limited. For example, a comprehensive testing of DFT and wave function methods by Zheng and Truhlar[7] against the DBH24/08 database of benchmark reaction energies[5,6] revealed that the performance of the best empirically-tailored hybrid DFT model chemistries approaches the 1 kcal/mol (4.2 kJ/mol) mean unsigned error, whereas the majority of hybrid DFT functionals result in average errors of several kilocalories per mole. The high-end many-body wave function methods appear to be the only way at the moment to reliably reduce the errors below the chemical-accuracy mark; e.g., the "gold standard" coupled-cluster singles, doubles, and perturbative triples-[CCSD(T)] method with a relatively large aug-cc-pCV(T+d)Z basis, while orders of magnitude more expensive than the DFT counterparts, was in error by only ~ 0.5 kcal/mol.

We should note that the electronic components of the molecular energies are the principal—but not the only—stumbling block here. The most accurate first-principles thermochemical models[124] inform us that prediction of relative energies to sub-1-kcal/mol accuracy reliably requires accounting for the effects of (1) anharmonicity, (2) special relativity (kinematic, Darwin, and spin-orbit terms), and (3) adiabatic coupling of electronic and nuclear motions. Yet the biggest remaining challenge is accurate computation of relative "bare" electronic energies: it requires (4) high-end wave function models and (5) reduction of the basis set error (typically, via basis set extrapolation). This chapter is concerned with the last issue, the basis set error, its role in prediction of the electronic reaction barriers, and to what extent it is alleviated by the use of novel explicitly correlated wave function methods.

The most troublesome origin of the basis set error of many-body wave functions is the sole use of Slater determinants as the building blocks. Standard wave function methods, such as CCSD(T) or multireference configuration interaction (MRCI), predict electron correlation energies with basis sets errors that decrease in atoms as $\mathcal{O}[(L_{\max}+1)^{-3}]$ for a basis set saturated to the angular momentum $L_{\max}$.[8] The basis set problem is due to the orbital product nature of Slater determinants; the Coulomb hole[36] that appears at short inter-electronic distances cannot be efficiently described by orbital products alone.[34] Explicitly correlated methods overcome this problem by modeling the Coulomb hole in terms that depend on the inter-electronic distances (r_{ij}). For atoms, the explicitly correlated methods of F12 type (pioneered by Kutzelnigg[9] and commonly known as "R12 methods") have basis set errors of $\mathcal{O}[(L_{\max}+1)^{-7}]$. The need to compute expensive three- and four-electron integrals, characteristic of all explicitly correlated methods, is avoided in F12 methods by the resolution of the identity;[64–66,68] only two-electron, albeit nonstandard, integrals are needed in F12 methods.

Rapid improvement of the F12 technology over the last decade has made it ready for use by nonspecialists. The modern F12 methods are efficient and competitive with extrapolation: to achieve the same basis set error, the approximate CC-F12 methods require a basis set of two cardinal numbers lower than the comparable standard CC computation. Despite the tremendous progress, the F12 technology is not yet sufficiently robust for black-box applications across the periodic table. One of the significant challenges, among many, is the requirement for particular orbital basis set types (OBS). Specifically, for robust performance in the F12 context, the OBS needs to include diffuse atomic orbitals; this requirement makes the applications of F12 methods difficult to large systems as well as to some electronically excited states. The only recommended options for F12 computations remain the cc-pVXZ-F12 series of F12-optimized basis sets of Peterson et al.[15–17] and the standard aug-cc-pVXZ series of Dunning et al.[18–20] In this work, our goal is to systematically evaluate these two series of basis sets, with an eye toward future improvement along the lines of the recent work of some of us.[125]

In this chaper, we document the extension of the perturbative coupled-cluster F12 method, CCSD(T)$_{\overline{\text{F12}}}$, developed in our group for the treatment of high-spin open-shell molecules and implemented in the open-source freely available Massively Parallel Quantum Chemistry

Table 4.1: The reactions including forward and reverse directions in the DBH24/08 database.[5–7,a]

Reaction No.	Reaction
	Heavy-Atom Transfer
1	$H + N_2O \leftrightarrow OH + N_2$
2	$H + ClH \leftrightarrow HCl + H$
3	$CH_3 + FCl \leftrightarrow CH_3F + Cl$
	Nucleophilic Substitution
4	$Cl^- \cdots CH_3Cl \leftrightarrow ClCH_3 \cdots Cl^-$
5	$F^- \cdots CH_3Cl \leftrightarrow FCH_3 \cdots Cl^-$
6	$OH^- + CH_3F \leftrightarrow HOCH_3 + F^-$
	Unimolecular and Association
7	$H + N_2 \leftrightarrow HN_2$
8	$H + C_2H_4 \leftrightarrow CH_3CH_2$
9	$HCN \leftrightarrow HNC$
	Hydrogen Transfer
10	$OH + CH_4 \leftrightarrow CH_3 + H_2O$
11	$H + OH \leftrightarrow O + H_2$
12	$H + H_2S \leftrightarrow H_2 + HS$

[a]There are 22 unique reaction barriers (the forward and reverse reactions in reaction 2 and 4 are the same).

(MPQC) package.[101] We assessed its performance for accurate studies of chemical reactivity by performing benchmark calculations of reaction barrier heights and thermochemical properties including electronic reaction energies, atomization energies, and enthalpies of formation. The performance of the CCSD(T)$_{\overline{F12}}$ method for reaction barrier heights is assessed by benchmarking against the DBH24/08 database (Table 4.1). The performance for electronic reaction energies is gauged for the HJO12 set of 12 isogyric reactions[1] (Table 4.2), which range from nearly thermoneutral to highly exothermic and involve breaking and forming an assortment of chemical bond. The performance for computation of the atomization energies and enthalpies of formation of the molecules is measured for the standard high-accuracy extrapolated ab initio thermochemistry (HEAT) set.[126] The HEAT test set includes 31 atoms and molecules for which both experimentally derived and theoretical enthalpies of formation (at 0 K) are available in better than a 1 kJ/mol agreement.

We performed two types of analyses targeting the two distinct uses of explicitly correlated CCSD(T) models: as a replacement for basis-set-extrapolated CCSD(T) in highly accurate composite methods like HEAT and as a distinct model chemistry for standalone applications.

Table 4.2: The HJO12 set of isogyric reactions involving 20 small molecules.

Reaction No.	Reaction
1	$CO+H_2 \rightarrow CH_2O$
2	$N_2+3H_2 \rightarrow 2NH_3$
3	$C_2H_2 + H_2 \rightarrow C_2H_4$
4	$CO_2 +4H_2 \rightarrow CH_4 +2H_2O$
5	$CH_2O+2H_2 \rightarrow CH_4 +H_2O$
6	$CO+3H_2 \rightarrow CH_4+H_2O$
7	$HCN+3H_2 \rightarrow CH_4+NH_3$
8	$HNO+2H_2 \rightarrow NH_3+H_2O$
9	$C_2H_2 + 3H_2 \rightarrow 2CH_4$
10	$CH_2 + H_2 \rightarrow CH_4$
11	$F_2 + H_2 \rightarrow 2HF$
12	$2CH_2 \rightarrow C_2H_4$

Hence, we analyzed in detail (1) the *basis set error* of each component of the CCSD(T)$_{\overline{\text{F12}}}$ contribution to the chemical energy difference in question and (2) the *total error* of the CCSD(T)$_{\overline{\text{F12}}}$ model chemistry relative to the benchmark values.

We begin with a brief description of the theoretical approach and computational details in Section 4.2. Results of computations and their discussion are given in Section 4.3. We summarize our findings and discuss their relevance to the future applications of F12 methods in Section 4.4.

4.2 Computational Methods

The original diagonal orbital-invariant formulation of the perturbative coupled-cluster F12 method, CCSD(T)$_{\overline{\text{F12}}}$, was reported for closed-shell species in Ref. 98. An open-shell variant of the CCSD(T)$_{\overline{\text{F12}}}$ method has been reported by our group[109] before based on the nondiagonal ansatz with optimized geminal coefficients. In this work, we implemented the open-shell method using the diagonal orbital-invariant (SP) ansatz of Ten-no;[68] this development is similar to work by others[96,127] in the context of the iterative CC-F12 methods. Thus, we only report the essential details of the open-shell variant of the method; the full programmable equations as implemented in the MPQC package are reported in the Supporting Information (Section 4.5).

4.2.1 The CCSD(T)$_{\overline{\text{F12}}}$ Formalism

The CCSD wave function captures the effects of two-electron correlations and the resulting orbital relaxation in a finite basis:

$$\Psi_{\text{CCSD}} \equiv \exp(\hat{T})|\Psi_0\rangle \tag{4.1}$$

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 \tag{4.2}$$

$$\hat{T}_1 \equiv t_a^i \tilde{a}_i^a \tag{4.3}$$

$$\hat{T}_2 \equiv \frac{1}{(2!)^2} t_{ab}^{ij} \tilde{a}_{ij}^{ab} \tag{4.4}$$

where Ψ_0 is the Hartree-Fock reference function (the notation is explained in the Appendix A). In the CCSD(T)$_{\overline{\text{F12}}}$ method, the CCSD wave function is extended perturbatively to account for the effect of three-electron correlations in the finite basis (via standard triple excitations), residual orbital relaxation effects (via standard single excitations into the complementary auxiliary basis set (CABS)), as well as for short-distance two-electron correlations (via explicitly correlated geminals).[109] The corresponding CCSD(T)$_{\overline{\text{F12}}}$ energy has four contributions:

$$E_{\text{CCSD(T)}_{\overline{\text{F12}}}} = E_{\text{CCSD}} + E_{(\text{T})} + E_{(2)_{\text{S}}} + E_{(2)_{\overline{\text{F12}}}}, \tag{4.5}$$

where the first two terms add up to the standard CCSD(T) energy, the next contribution is the usual "CABS singles" correction,[80] and the last contribution is due to the explicit correlated F12 terms. In the diagonal orbital-invariant approach to CCSD(T)$_{\overline{\text{F12}}}$,[98] the $E_{(2)_{\overline{\text{F12}}}}$ contribution is evaluated directly from the usual Hylleraas functional, without any parameter optimization:

$$E_{(2)_{\overline{\text{F12}}}} = \langle 1|\hat{H}^{(0)}|1\rangle + \langle 0|\hat{H}^{(1)}|1\rangle + \langle 1|\hat{H}^{(1)}|0\rangle. \tag{4.6}$$

The zeroth- and first-order Hamiltonians are defined from the similarity-transformed CCSD Hamiltonian, $\bar{H} \equiv \exp(-\hat{T})\hat{H}\exp(\hat{T})$, by matrix (Löwdin) partitioning[108] so that $\hat{H}^{(0)}$ in Eq. 4.6 becomes the standard Fock operator, $\hat{F}$, and $\hat{H}^{(1)}$ becomes $\bar{H}$. The left- and right-hand zeroth-order wave functions are the standard left- and right-hand CCSD "wave functions":

$$|0\rangle \equiv |\Psi_0\rangle, \tag{4.7}$$

$$\langle 0| \equiv \langle \Psi_0|(1 + \hat{\Lambda}), \tag{4.8}$$

where Λ represents the undetermined multipliers of the CCSD Lagrangian.[1] Since $\hat{\Lambda}$ is not necessary for computing the CCSD energy (it is, however, necessary for computing molecular properties), by analogy with the CCSD(T) method we invoke approximation $\hat{\Lambda} \approx \hat{T}^{\dagger}$.[107]

The first-order wave function,

$$|1\rangle \equiv \frac{1}{2!}\sum_{ij} |\Gamma_{ij}^{ij}\rangle, \tag{4.9}$$

is composed of geminal functions, defined as

$$|\Gamma_{ij}^{ij}\rangle = \frac{1}{2!}\bar{R}_{\alpha\beta}^{ij}\tilde{\alpha}_{ij}^{\alpha\beta}|\Psi_0\rangle, \tag{4.10}$$

where $\bar{R}_{\alpha\beta}^{ij}$ is the antisymmetrized matrix element of the geminal correlation factor $f(r_{12})$:

$$R_{\alpha\beta}^{ij} = \langle\alpha\beta|\hat{Q}_{12}f(r_{12})|ij\rangle. \tag{4.11}$$

Projector $\hat{Q}_{12}$ ensures that the geminal functions in Eq. 4.10 are strongly orthogonal (orthogonal to the $|\Psi_0\rangle$ and to the single excitations) as well as orthogonal to the standard double excitations:[66]

$$\hat{Q}_{12} = (1-\hat{O}_1)(1-\hat{O}_2) - \hat{V}_1\hat{V}_2, \tag{4.12}$$

where $\hat{O}$ and $\hat{V}$ are the projectors on the occupied and virtual orbitals.

The geminal correlation factor is designed so that functions $|\Gamma_{ij}^{ij}\rangle$ describe the short-range two-electron correlation according to the first-order cusp condition for the pair function $|\psi_{ij}\rangle$:

$$r_{12} \to 0: \quad |\psi_{ij}^{(1)}\rangle = C_l r_{12}|\psi_{ij}^{(0)}\rangle, \tag{4.13}$$

where C_l is the cusp coefficient that depends on the permutational symmetry of $\psi_{ij}^{(0)}$. In practice, it is sufficient to consider only two cases: for singlet pairs $C_0 = 1/2$ and for triplet pairs $C_1 = 1/4$ (more exotic cases, such as the unnatural parity singlet for which $C_2 = 1/6$, can be realized in heavy atoms[8]). The appropriate selection of the proper coefficient can be attained using a spatial coordinate permutation operator[68] or, equivalently, with spin projectors:

$$\hat{P}_0 = |\alpha\beta\rangle_0\langle\alpha\beta|_0 \tag{4.14}$$

$$\hat{P}_1 = |\beta\beta\rangle\langle\beta\beta| + |\alpha\beta\rangle_1\langle\alpha\beta|_1 + |\alpha\alpha\rangle\langle\alpha\alpha| \tag{4.15}$$

with

$$|\alpha\beta\rangle_0 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\}, \tag{4.16}$$

$$|\alpha\beta\rangle_1 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) + \beta(1)\alpha(2)\}. \tag{4.17}$$

Thus, the spin-adapted geminal correlation factor is written as

$$f(r_{12}) \equiv (C_0\hat{P}_0 + C_1\hat{P}_1)\gamma(r_{12}). \tag{4.18}$$

In this work, $\gamma(r_{12})$ was the standard exponential correlation factor of Ten-no[71] expanded as a linear combination of $N_g = 6$ Gaussian geminals:[73]

$$\gamma(r_{12}) \equiv -l_c\exp(-r_{12}/l_c) \approx \sum_i^{N_g} c_i\exp(-\alpha_i r_{12}^2). \tag{4.19}$$

The correlation length scale, $l_c \sim 1\,a_0$, was set to a value that depends empirically on the orbital basis set.

With the new form of geminal functions, the second-order F12 correction (Eq. 4.6) for the first-order wave function ansatz (Eq. 4.9) has the following form:

$$E_{(2)_{\overline{\text{F12}}}} = \sum_{i<j} \varepsilon_{ij}^{(2)}, \tag{4.20}$$

$$\varepsilon_{ij}^{(2)} = 2\tilde{V}_{ij}^{ij} + \tilde{B}_{ij}^{ij}, \tag{4.21}$$

where $\tilde{V}_{ij}^{ij}$ is the geminal matrix of the first-order Hamiltonian and has the following form:

$$\tilde{V}_{ij}^{ij} \equiv V_{ij}^{ij} + \frac{1}{2}(V_{ab}^{ij} + C_{ab}^{ij})t_{ij}^{ab} + V_{ia}^{ij}t_j^a + V_{aj}^{ij}t_i^a, \tag{4.22}$$

$$V_{pq}^{ij} \equiv \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{g}_{pq}^{\alpha\beta}, \tag{4.23}$$

$$C_{ab}^{ij} \equiv F_a^\alpha \bar{R}_{\alpha b}^{ij} + F_b^\alpha \bar{R}_{a\alpha}^{ij}. \tag{4.24}$$

$\tilde{B}_{ij}^{ij}$ is the geminal matrix of the zero-order Hamiltonian. If we assume canonical Hartree-Fock orbitals, it is expressed as

$$\tilde{B}_{ij}^{ij} \equiv B_{ij}^{ij} - (F_i^i + F_j^j)X_{ij}^{ij}, \tag{4.25}$$

$$B_{ij}^{ij} \equiv \bar{R}_{\alpha\beta}^{ij}F_\gamma^\beta \bar{R}_{ij}^{\alpha\gamma}, \tag{4.26}$$

$$X_{ij}^{ij} \equiv \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{R}_{ij}^{\alpha\beta}. \tag{4.27}$$

The infinite-range sums in the matrices V, C, and X were approximated by finite sums over the complementary auxiliary basis set (CABS),[66] while B used approximation C of Kedzuch et al.[67] as well as the CABS approach (complete details have been described by one of us elsewhere[120]).

Because the (T) correction to CCSD(T) is most commonly formulated in the spin-unrestricted framework, we assumed that the spatial parts of spin-orbitals are spin-dependent. After integrating out the spin degrees of freedom, we arrived at the final expression for Eq. 4.20:

$$E_{(2)_{\overline{\text{F12}}}} = \sum_{I<J} \varepsilon_{IJ}^{(2)} + \sum_{I,\bar{J}} \varepsilon_{I\bar{J}}^{(2)} + \sum_{\bar{I}<\bar{J}} \varepsilon_{\bar{I}\bar{J}}^{(2)}, \tag{4.28}$$

$$\varepsilon_{IJ}^{(2)} = 2C_1\tilde{V}_{IJ}^{IJ} + C_1^2\tilde{B}_{IJ}^{IJ}, \tag{4.29}$$

$$\begin{aligned}\varepsilon_{I\bar{J}}^{(2)} = {} & (C_0 + C_1)\tilde{V}_{I\bar{J}}^{I\bar{J}} + (C_0 - C_1)\tilde{V}_{I\bar{J}}^{\bar{J}I} + \frac{(C_0 + C_1)^2}{4}\tilde{B}_{I\bar{J}}^{I\bar{J}} \\ & + \frac{C_0^2 - C_1^2}{4}\tilde{B}_{I\bar{J}}^{\bar{J}I} + \frac{C_0^2 - C_1^2}{4}\tilde{B}_{\bar{J}I}^{I\bar{J}} + \frac{(C_0 - C_1)^2}{4}\tilde{B}_{\bar{J}I}^{\bar{J}I},\end{aligned} \tag{4.30}$$

$$\varepsilon_{\bar{I}\bar{J}}^{(2)} = 2C_1V_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} + C_1^2\tilde{B}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}}. \tag{4.31}$$

The detailed expressions for the spin-free matrices $\tilde{V}$ and $\tilde{B}$ are presented in the Supporting Information (Section 4.5).

4.2.2 Computational Details

The (U)QCISD/MG3 molecular geometries for the DBH24/08 database were obtained from Ref. 5. Two types of orbital basis sets were utilized for the reaction barrier calculation:

- The cc-pVXZ-F12 family of basis sets (X = D, T, Q) designed by Peterson and co-workers[15] to be used specifically with explicitly correlated methods. The complementary auxiliary basis set (CABS) was constructed from the matching cc-pVXZ-F12/OptRI basis sets[16] using the CABS+ approach.[66]
- The standard aug-cc-pVXZ family of basis sets (X = D, T, Q) that was designed by Dunning et al.[18–20] for computations with standard correlated methods (for S and Cl atoms, the corresponding basis sets augmented with tight d functions, aug-cc-pV(X+d)Z,[128] were used). The complementary auxiliary basis set (CABS) was constructed using the matching aug-cc-pVXZ/OptRI basis sets[84] using the CABS+ approach.[66]

The evaluation of Eq. 4.28 utilized the robust density fitting[118] in a cc-pV(X+1)Z-RI basis set for the cc-pVXZ-F12 calculation and an aug-cc-pVXZ-RI basis set for the aug-cc-pVXZ calculation, where X is the cardinal number of the corresponding orbital basis; the conventional CC wave functions and energies were evaluated without the density fitting. The recommended correlation length scales[15] were used; correlation factors were expanded in terms of six Gaussian geminals.[3]

Although there are only 22 unique barriers in the DBH24/08 database, to be consistent with Zheng et al.[7] in our analysis we also considered all 24 barriers statistically independent. The recommended spin-orbit energies[7] were used to correct the energies of Cl, O, OH, and HS radicals. Other relativistic effects were neglected. In the calculations of reaction barriers here, only the valence electron correlations were considered (core correlation contributions were estimated to be below 0.1 kcal/mol, as described below).

Experimental geometries were used for the 20 closed-shell molecules involved in the HJO12 set of reactions.[1] The computational procedure was identical to that used for the DBH24/08 set. The core-correlation corrections were obtained from Ref. 1.

The CCSD(T)/cc-pVQZ geometries from Ref. 126 were used for the molecules in the HEAT test set. Two core correlation basis sets were utilized for the calculations of atomization energies and enthalpies of formation. The first sets are the cc-pCVXZ-F12 (X = D, T, Q) basis sets, which are also designed by Peterson et al.[83] for the explicitly correlated methods, and the second sets are the standard aug-cc-pCVXZ (X = D, T, Q) basis set,[18,20] which are used for the standard correlation methods. The CABS was constructed with the cc-pCVXZ-F12/OptRI basis sets[83] using the CABS+ approach.[66] The aug-cc-pwCV(X+1)Z-RI basis set was used for the robust density fitting[118] in the evaluation of Eq. 4.28. The anharmonic

zero-point energies calculated within the HEAT model were used, while other relativistic corrections were neglected.

All explicitly correlated calculations were performed using the trunk version of the Massively Parallel Quantum Chemistry (MPQC) package[101] (it can be obtained for free under the GNU General Public License (GPL) at http://www.mpqc.org/). The CC wave functions and energies were computed with the PSI3 suite.[129]

4.3 Discussion of Results

4.3.1 Basis Set Errors of Reaction Barriers

The mean unsigned basis set errors of the standard and explicitly correlated CCSD(T) energies (and their selected components) for the DBH24/08 reaction barriers are presented in Figure 4.1. The basis set errors (BSE) were defined with respect to the complete basis set (CBS) limits defined as follows: (1) the HF(2)$_{\text{S}}$/cc-pVQZ-F12 energies were used as the CBS Hartree-Fock limit; (2) the CBS limits for the CCSD and (T) correlation energy components were obtained using the standard X^{-3} extrapolation formula[33] from the aug-cc-pVTZ and aug-cc-pVQZ values.

First, we note that the basis set errors of the Hartree-Fock contributions to reaction barriers cannot be neglected ($\sim$ 4 kJ/mol with the aug-cc-pVDZ basis set). Although the HF BSEs are smaller with the cc-pVXZ-F12 basis sets than with the aug-cc-pVXZ counterparts (due to the extra s and p functions in the former), at the double-ζ level the use of CABS singles correction is mandatory. Even with the correction, the aug-cc-pVDZ HF energy is still in error by $\sim$ 1 kJ/mol.

The basis set errors of CCSD energies are large: $\sim$ 3-4 kJ/mol with the double-ζ basis set, and $\sim$ 1.5 kJ/mol with a triple-ζ basis set. Explicitly correlated CCSD energies have greatly reduced basis set errors: for example, the mean unsigned BSE of CCSD(2)$_{\overline{\text{F12}}}$/cc-pVDZ-F12 is 1.38 kJ/mol. Note that when used with non-F12 methods the F12-optimized cc-pVXZ-F12 basis sets result in larger basis set errors than the standard aug-cc-pVXZ basis sets, as expected. In combination with F12 methods, only the cc-pVDZ-F12 basis set is clearly preferred over the aug-cc-pVDZ counterpart; at the triple- and quadruple-ζ levels, the two basis set families result in essentially identical BSEs, on the order of 0.2 kJ/mol or below. Unfortunately, it is not possible to judge the relative performance of the two families any further because of the limited accuracy of the CBS limits. We estimate the current CBS limit for CCSD energy obtained by the extrapolation to be only accurate to 0.1 kcal/mol ($\sim$ 0.4 kJ/mol). This figure is in line with the mean unsigned difference of 0.25 kJ/mol between the cc-pVQZ-F12 and aug-cc-pVQZ CCSD(2)$_{\overline{\text{F12}}}$ energies. It is unlikely that the CBS limits can be deduced more accurately without performing time-consuming quintuple-ζ computations; however, since the 0.1 kcal/mol accuracy is quite satisfactory, we do not

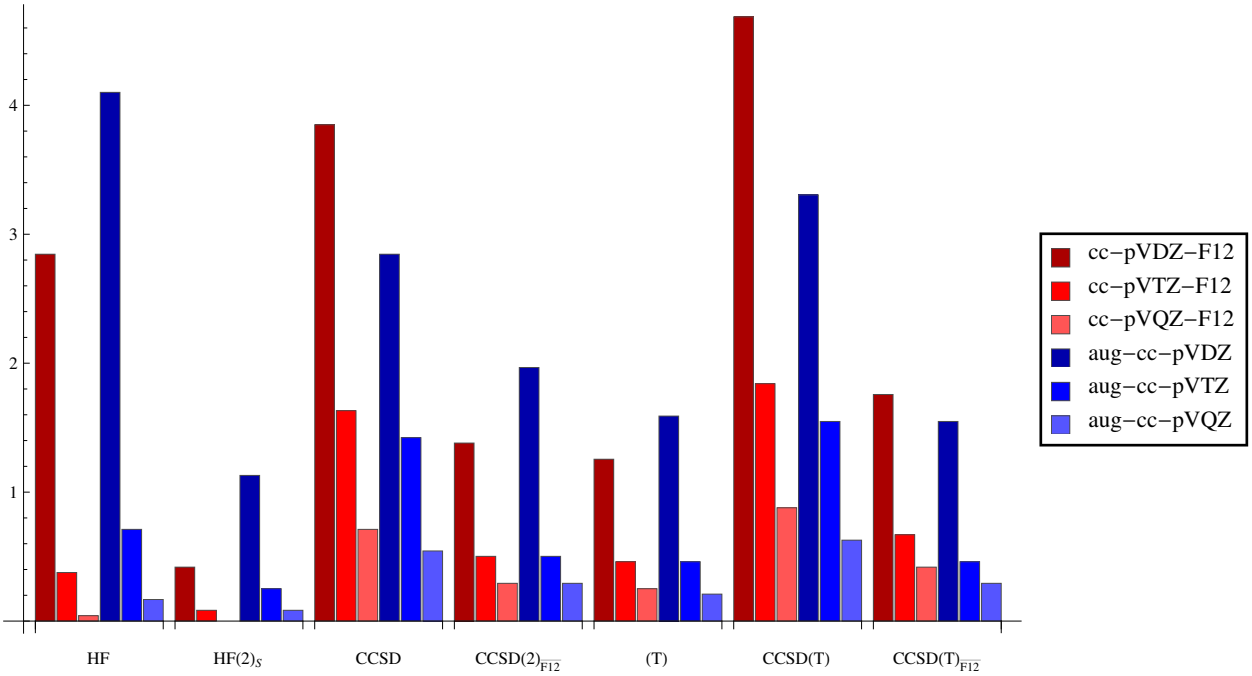

Figure 4.1: Average unsigned basis set errors (kJ/mol) of various electronic contributions to the reaction barriers. "HF" and "HF(2)$_S$" refer to the Hartree-Fock energies without and with inclusions of the CABS singles correction (Eq. 4.5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

consider it worthwhile at the moment.

Finally, the CCSD(T) BSE is greatly reduced also by the use of the explicit correlation. Only a double-ζ basis set is sufficient with the CCSD(T)$_{\overline{\text{F12}}}$ method for $\sim$ 1.6 kJ/mol BSE; triple-ζ basis sets further reduce basis set errors to $\sim$ 0.5 kJ/mol. The conventional CCSD(T) method requires basis sets one cardinal number greater (triple and quadruple-ζ, respectively) to achieve the same precision. Even with the explicitly correlated CCSD method, the remaining basis set errors are comparable to the basis set errors of the (T) contribution to the CCSD(T) energy. Therefore it is, in our opinion, acceptable to treat the (T) energy without the use of explicitly correlated three-body functions, unless extremely high accuracy is needed.

We observed a peculiar behavior of basis set errors of CCSD(2)$_{\overline{\text{F12}}}$ and (T) contributions: (1) with the cc-pVXZ-F12 basis sets, the two errors add up constructively; thus the mean unsigned error of CCSD(T)$_{\overline{\text{F12}}}$ is greater than that of CCSD(2)$_{\overline{\text{F12}}}$, e.g., for cc-pVDZ-F12: 1.76 kJ/mol vs. 1.38 kJ/mol. (2) With the aug-cc-pVXZ basis sets, the errors cancel; thus, the mean unsigned error of CCSD(T)$_{\overline{\text{F12}}}$ is smaller than that of CCSD(2)$_{\overline{\text{F12}}}$, e.g., for aug-cc-pVDZ: 1.55 kJ/mol vs. 1.97 kJ/mol. It is not clear whether this behavior is systematic;

it will be important in the discussion of total errors in Section 4.3.5.

4.3.2 Basis Set Errors of Reaction Energies

Figure 4.2 shows the basis set errors of the HF energies and various correlation contributions to the electronic reaction energies for the HJO12 set[1] of isogyric reactions. The CBS limits for the basis set errors were defined as follows: (1) the CBS Hartree-Fock values were the reaction energies computed with the aug-cc-pV6Z basis set. (2) The valence (T) CBS limit was obtained by the basis set extrapolation with the Schwenke method[130] using aug-cc-pVQZ and aug-cc-pV5Z basis sets. (3) The CBS limit for the CCSD(T) valence correlation energies was obtained by subtracting from the CBS CCSD(T) energies[1] (obtained by the X^{-3} extrapolation using the cc-pCV5Z and cc-pCV6Z basis sets) the cc-pV6Z HF energies and the core-correlation energies given in Ref. 1. (4) The CCSD CBS limit was obtained by subtracting the Schwenke extrapolated CBS limit of the valence (T) energies from the CCSD(T) CBS limit.

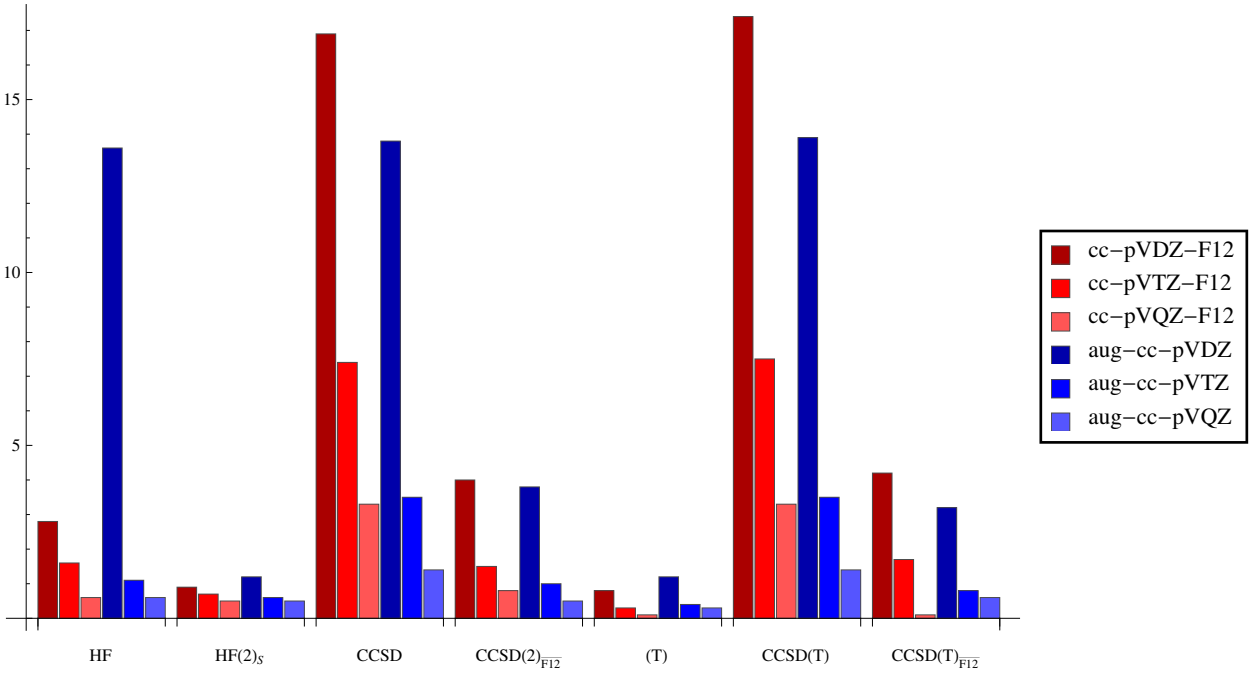

Figure 4.2: Average unsigned basis set errors (kJ/mol) of various electronic contributions to the reaction energies. "HF" and "HF(2)$_S$" refer to the Hartree-Fock energies without and with inclusions of the CABS singles correction (Eq. 4.5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

A very large HF basis set error results from the use of the aug-cc-pVDZ basis set (13.6 kJ/mol). Fortunately, the CABS singles correction reduces the basis set error to 1.2 kJ/mol. Although the HF BSE to reaction energies with the cc-pVDZ-F12 basis set is much smaller (2.8 kJ/mol), the CABS singles correction is still needed to reduce the error around 1 kJ/mol. At the triple-ζ level, however, the HF BSE with the F12 basis set is slightly larger than that with the non-F12 counterpart, but the HF BSE with the two basis sets become identical at the quadruple-ζ level.

The basis set errors of CCSD energies are large, especially with the double-ζ basis sets (over 13 kJ/mol). Explicitly correlated CCSD energies have much smaller basis set errors (~ -4 kJ/mol for X = D and ~ 1 kJ/mol for X = T, Q). Compared to the CCSD(2)$_{\overline{\text{F12}}}$ BSEs, the (T) BSEs are noticeably smaller, especially at the double-ζ level. Unexpectedly, we note that the cc-pVXZ-F12 (X = D, T) basis sets give slightly larger CCSD(2)$_{\overline{\text{F12}}}$ BSE than the corresponding aug-cc-pVXZ basis sets for all of the cases but have slightly smaller basis set errors for the (T) contributions than the aug-cc-pVXZ basis sets.

The basis set errors of the standard CCSD(T) energies are very large and slowly convergent, but the use of the explicit correlation significantly reduces the basis set errors. The triple-ζ CCSD(T)$_{\overline{\text{F12}}}$ calculations have BSEs twice smaller than the quadruple-ζ CCSD(T) calculations. With the quadruple-ζ basis sets, the basis set errors decrease to under 1 kJ/mol. Overall, the BSEs from the two basis set families do not differ significantly. While the aug-cc-pVXZ basis sets yield smaller BSEs than the cc-pVXZ-F12 basis sets when X = D, T, the aug-cc-pVQZ basis set gives a larger error than the cc-pVQZ-F12 basis set, which has a BSE close to zero.

4.3.3 Basis Set Errors of Atomization Energies

The basis set errors of the various electronic contributions to the atomization energies of the HEAT test set are shown in Figure 4.3. The HF/aug-cc-pCV5Z energies were defined as the CBS HF limit; the CBS limits for the CCSD and (T) correlation energy components were obtained using the X^{-3} extrapolation formula[33] with the aug-cc-pCVQZ and aug-cc-pCV5Z values.

As expected, the F12 basis sets give smaller HF basis set errors than the non-F12 basis sets. The basis set errors of the double-ζ HF contributions are large (> 8 kJ/mol), and the use of the CABS singles correction is necessary to reduce the errors around 1 kJ/mol. At the triple- and quadruple-ζ levels, however, the HF BSEs become quite small; the use of the CABS singles correction can be optional.

The CCSD basis set errors are very large with the small basis sets (> 40 kJ/mol and > 19 kJ/mol for X = D and T, respectively). Even with the large quadruple-ζ basis sets, the BSEs are still around 8 kJ/mol. Unexpectedly, at the double- and triple-ζ levels, the F12 basis sets yield slightly smaller CCSD BSEs than the standard non-F12-optimized aug-cc-pCVXZ

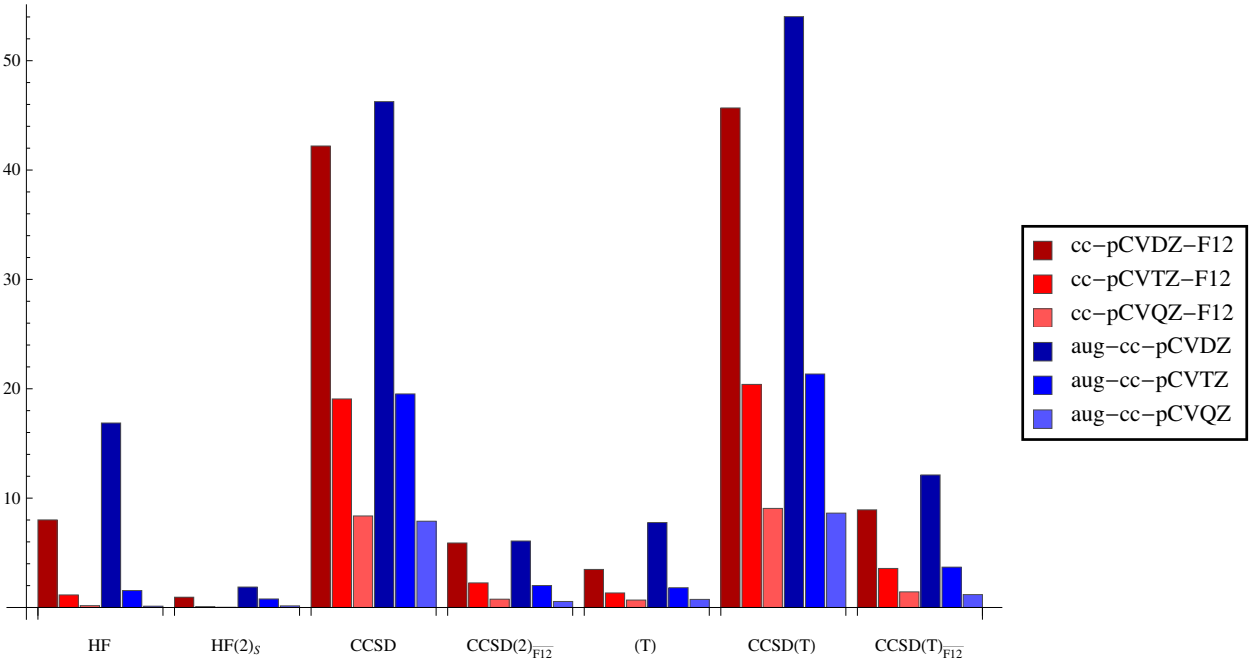

Figure 4.3: Average unsigned basis set errors (kJ/mol) of various electronic contributions to the atomization energies. "HF" and "HF(2)$_S$" refer to the Hartree-Fock energies without and with inclusions of the CABS singles correction (Eq. 4.5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in text.

basis sets. The CCSD BSEs are significantly reduced by the use of the explicit correlation. In fact, the double-ζ CCSD(2)$_{\overline{\text{F12}}}$ BSEs are noticeably smaller than the quadruple-ζ CCSD BSEs. With the quadruple-ζ basis sets, the BSEs of the CCSD(2)$_{\overline{\text{F12}}}$ correlation energies are reduced to below 1 kJ/mol. The two basis sets yield similar CCSD(2)$_{\overline{\text{F12}}}$ BSEs.

The (T) BSEs are slightly smaller than the explicitly correlated CCSD BSEs in most cases. The only exception is the (T) BSE with the aug-cc-pCVDZ basis set. As a result, while the BSE of the CCSD(T)$_{\overline{\text{F12}}}$ correlation energies with the aug-cc-pCVDZ basis set (12.12 kJ/mol) is noticably larger than the BSE of the CCSD(T) correlation energies with the aug-cc-pCVQZ basis set (8.63 kJ/mol), the CCSD(T)$_{\overline{\text{F12}}}$ BSE with the cc-pCVDZ-F12 basis set (8.93 kJ/mol) is smaller than the CCSD(T) BSE with the cc-pCVQZ-F12 basis set (9.06 kJ/mol), which follows the same pattern for the CCSD(2)$_{\overline{\text{F12}}}$ BSE. With the quadruple-ζ basis sets, the basis set errors of the explicitly correlated CCSD(T) correlation energies reduce to $\sim$ 1 kJ/mol. Different from the calculations in the previous sections, we find the aug-cc-pCVXZ calculations yield larger basis set errors than the cc-pCVXZ-F12 basis set when X = D and T, but give slightly smaller error than the corresponding F12 basis set when X = Q.

4.3.4 Basis Set Errors of Enthalpies of Formation

Figure 4.4 presents the basis set errors of the HF energies and various contributions to the enthalpies of formation. The CBS limits were defined in the same way as the previous section.

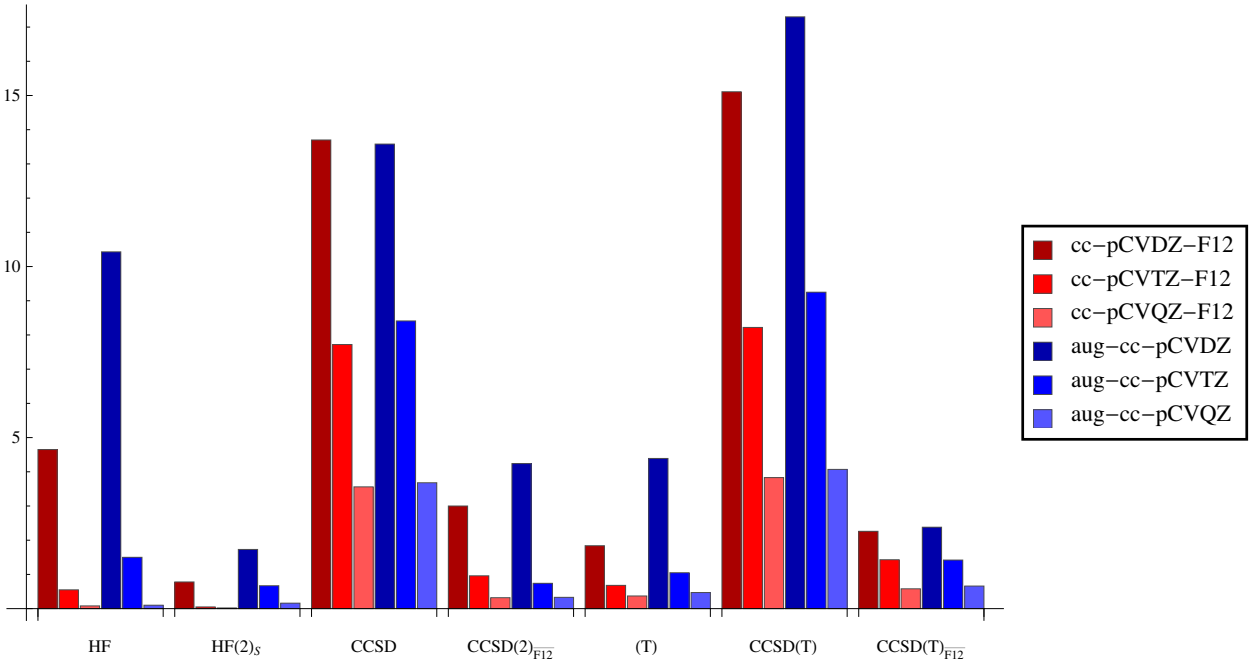

Figure 4.4: Average unsigned basis set errors (kJ/mol) of various electronic contributions to the enthalpies of formation. "HF" and "HF(2)$_S$" refer to the Hartree-Fock energies without and with inclusions of the CABS singles correction (Eq. 4.5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

Again, the HF basis set errors with the double-ζ basis sets are large, and the CABS singles correction reduces the errors to around 1∼2 kJ/mol. As we increase the size of the basis set, the use of the CABS singles correction becomes optional since the HF BSEs become small.

The basis set errors of the CCSD correlation energies range from $\sim$ 13 kJ/mol to $\sim$ 3 kJ/mol. The range is reduced to $\sim$ 4 kJ/mol to $\sim$ 0.3 kJ/mol with the explicit correlation. The two basis set families generate close basis set errors, and neither of them shows consistently better performance than the other. Overall, the CCSD(T)$_{\overline{\text{F12}}}$ BSEs are close to the (T) BSEs.

Again, the explicitly correlation significantly improves the basis convergence of the CCSD(T) correlation energies. The CCSD(T)$_{\overline{\text{F12}}}$ BSEs with the double-ζ basis sets are almost twice smaller than the CCSD(T) BSEs with the quadruple-ζ basis sets. With the large quadruple-ζ

Table 4.3: The mean unsigned errors and maximum unsigned errors (kJ/mol) of DBH24/08 reaction barriers computed with CCSD and CCSD(T) methods using the cc-pVXZ-F12 and aug-cc-pVXZ basis set families.[a]

Method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
			cc-pVXZ-F12			
CCSD	11.89	9.84	9.77	64.46	44.34	41.53
CCSD(2)$_{\overline{\text{F12}}}$	9.40	9.63	9.83	40.68	39.77	40.03
CCSD(T)	6.20	2.27	1.30	30.34	7.86	4.30
CCSD(T)$_{\overline{\text{F12}}}$	1.93	1.16	1.18	6.56	3.28	2.80
			aug-cc-pVXZ			
CCSD	7.59	9.14	9.63	47.62	43.06	41.20
CCSD(2)$_{\overline{\text{F12}}}$	8.04	9.36	9.76	32.85	38.52	38.95
CCSD(T)	5.21	2.37	1.24	16.06	6.99	4.12
CCSD(T)$_{\overline{\text{F12}}}$	1.72	1.05	1.08	4.44[b]	2.45	2.54[c]

[a]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 4.1). [b]Reverse Reaction 6. [c]Reverse Reaction 3.

basis sets, the basis set errors of explicitly correlated CCSD(T) correlation energies reduce to around 0.6 kJ/mol. These values are close to the corresponding overall errors of the HEAT model (0.24 kJ/mol). The BSEs from the calculations with the two basis sets are very close, although the F12 basis sets give slightly better results for the standard and explicitly correlated CCSD(T) correlation energies than the non-F12 basis sets.

4.3.5 Overall Performance of CCSD(T)$_{\overline{\text{F12}}}$ for Reaction Barriers

To access the overall performance of the CCSD(T)$_{\overline{\text{F12}}}$ model chemistry for computing reaction barriers, we computed the mean unsigned errors (MUEs) and maximum unsigned errors (MaxUEs) relative to the benchmark electronic barriers in the DBH24/08 database; the results are presented in Table 4.3.

As expected, the barriers computed with the CCSD method have large errors: the majority of the resulting MUEs are larger than 9 kJ/mol, and the MaxUEs are around or more than 41 kJ/mol. The inclusion of the F12 correction into the CCSD method reduces the MUEs only slightly. This should not be surprising: the F12 correction only reduces the basis set error and does not affect the method error.

We noticed that the MaxUEs are all from the reverse reaction barrier of reaction 1 (OH + $N_2 \rightarrow H + N_2O$), which has a very high reaction barrier (345.05 kJ/mol). To evaluate the

MaxUE influence on the MUEs, we also computed the MUEs of different methods without the MaxUE values. However, the new MUEs do not show a significant decrease; the errors are still on the same level as those including the MaxUEs.

The inclusion of the perturbative triple correction (T) is crucial to obtain accurate reaction barriers, as it becomes readily apparent from comparing the CCSD and CCSD(T) barriers. Yet a triple-ζ basis is needed to reduce the CCSD(T) MUEs to around 2 kJ/mol; the corresponding triple-ζ MaxUEs are still around or above 7 kJ/mol. Only with the large quadruple-ζ basis sets do the maximum errors of the CCSD(T) barriers reduce to $\sim$ 1 kJ/mol. However, for the largest DBH24/08 test cases, each increase of the cardinal number increases the computational cost of the CCSD(T) computation by an order of magnitude. Hence, with the transition from aug-cc-pVDZ to aug-cc-pVQZ the mean and maximum CCSD(T) errors reduce by a factor of 4, whereas the cost increases by a factor of $\mathcal{O}(100)$. The slow basis set convergence is clearly a severe problem for computing the reaction barriers in the DBH24/08 set and must be dealt with to unleash the inherent high accuracy of the CCSD(T) method.

Fortunately, the basis set errors of CCSD(T) are reduced dramatically by switching to the CCSD(T)$_{\overline{\text{F12}}}$ method. The use of F12 terms reduces the errors by an equivalent of 1 to 2 cardinal numbers (i.e., a double-ζ F12 result is between the accuracy of conventional triple- and quadruple-ζ energies). Only a double-ζ basis set is sufficient to reduce the mean basis set errors to below 2 kJ/mol. The mean errors reduce further to close to 1 kJ/mol with a triple-ζ basis; the use of a quadruple-ζ basis results in no additional error decrease. The maximum errors are also reduced dramatically compared to the standard CCSD(T) method, below 4 kJ/mol with triple-ζ basis sets. The mean and maximum errors obtained with the cc-pVXZ-F12 basis sets of Peterson and co-workers are greater than those obtained with the aug-cc-pVXZ counterparts, even though the former family was designed specifically for the use with F12 methods and have been found to outperform the standard basis sets.[85] The double-ζ results are especially surprising as the cc-pVDZ-F12 basis set is considerably larger than the aug-cc-pVDZ counterpart (the cc-pVDZ-F12 basis set has one more s and two more p functions). On the other hand, while the cc-pV{T,Q}Z-F12 basis sets are still larger than the corresponding aug-cc-pV{T,Q}Z basis sets for non-hydrogen atoms, they are actually smaller for hydrogen atom. For the hydrogen atom, the cc-pVTZ-F12 basis set has one less d function than the aug-cc-pVTZ basis set, and the cc-pVQZ-F12 basis set has one less d and f functions comparing to the aug-cc-pVQZ basis set.

To further investigate the effects of diffuse functions on the hydrogen atom, we carried the aug-cc-pVXZ calculations with the cc-pVXZ basis sets for hydrogen atoms. We find that the mean and maximum errors of the calculations (shown in Table 4.4) become larger, especially for the calculations at the double-ζ level. However, the mean errors of the triple- and quadruple-ζ level calculations are very close to those of the corresponding cc-pV{T,Q}Z-F12 calculations. This indicates that the extra diffuse functions are not the reason for the better performance of the aug-cc-pVXZ basis versus the cc-pVXZ-F12 basis. The discussion of basis set errors in Section 4.3.1 suggests that the smaller overall errors obtained with the aug-cc-pVXZ basis sets are perhaps due to a fortuitous cancellation of the CCSD(2)$_{\overline{\text{F12}}}$

Table 4.4: The CCSD(T) and CCSD(T)$_{\overline{\text{F12}}}$ mean and maximum unsigned errors (kJ/mol) of DBH24/08 reaction barriers using the aug-cc-pVXZ basis sets for non-hydrogen atoms and the cc-pVXZ basis sets for hydrogen atom.[a]

Method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
CCSD(T)	6.67	2.76	1.37	17.90	7.40	4.26
CCSD(T)$_{\overline{\text{F12}}}$	2.31	1.20	1.12	5.09[b]	3.54[c]	2.84[c]

[a]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 4.1). [b]Forward Reaction 12. [c]Reverse Reaction 12.

Table 4.5: The CCSD(T)$_{\overline{\text{F12}}}$ mean and maximum unsigned errors (kJ/mol) of DBH24/08 reaction barriers with the cc-pVXZ-F12 and aug-cc-pVXZ basis set families using different geminal exponents (β).[a,b]

Basis	MUE			MaxUE		
	$\beta = 0.8$	$\beta = 1.0$	$\beta = 1.2$	$\beta = 0.8$	$\beta = 1.0$	$\beta = 1.2$
cc-pVDZ-F12	2.08	1.91	2.07	8.14	5.67	5.22
cc-pVTZ-F12	1.10	1.16	1.22	3.63	3.28	3.24
aug-cc-pVDZ	1.65	1.65	1.81	4.17	4.27[c]	4.36[c]
aug-cc-pVTZ	1.08	1.05	1.05	2.59	2.37[d]	2.45

[a]The optimal geminal exponents for the aug-cc-pVDZ, aug-cc-pVTZ, cc-pVDZ-F12, and cc-pVTZ-F12 basis sets are 1.1, 1.2, 0.9, and 1.0, respectively.[15] [b]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 4.1). [c]Reverse Reaction 6. [d]Reverse Reaction 12.

and (T) energies. This, at the very least, suggests that the cc-pVXZ-F12 basis sets may need a more thorough benchmarking for a variety of applications before the use of standard non-F12-optimized basis sets is eliminated.

In addition, we performed the cc-pVXZ-F12 and aug-cc-pVXZ (X = D, T) calculations with different geminal exponents (0.8, 1.0 and 1.2). The results are listed in Table 4.5. As expected, the calculations at the triple-ζ level using different geminal exponents give very similar results, which is also the case for the aug-cc-pVDZ calculations. For the cc-pVDZ-F12 calculations, while the MUEs of the calculations with different geminal exponents are still close, the discrepancy between the corresponding MaxUEs is much larger. We also found the calculations with optimal geminal exponents do not necessarily give the best results. This suggests that the choice of geminal exponent is not crucial to the quality of the calculations as long as the geminal exponent is chosen in a reasonable range.

In the calculations above, we only included the valence electron correlations. To address

Table 4.6: The mean unsigned errors and maximum unsigned errors (kJ/mol) of the electronic reaction energies for the HJO12 isogyric reactions.[1, a]

Method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
			cc-pVXZ-F12			
CCSD	11.5	6.4	6.8	37.5	24.5	19.7
CCSD(2)$_{\overline{\text{F12}}}$	6.0	7.5	8.3	15.3	16.2	16.1
CCSD(T)	15.3	8.8	3.7	32.8[b]	17.5[b]	8.4[b]
CCSD(T)$_{\overline{\text{F12}}}$	4.7	2.3	1.6	13.7[b]	6.4[b]	4.3[b]
			aug-cc-pVXZ			
CCSD	15.4	9.7	9.5	45.9	27.3	20.1
CCSD(2)$_{\overline{\text{F12}}}$	6.8	8.3	9.0	14.6	16.6	16.4[c]
CCSD(T)	10.6	3.7	1.8	35.1	13.3	5.4
CCSD(T)$_{\overline{\text{F12}}}$	3.7	1.5	1.2	7.5[b]	3.6[b]	3.1[c]

[a]Unless noted specifically, the maximum errors were observed for the Reaction 12 (see Table 4.2). [b] Reaction 2. [c] Reaction 11.

the influence of the core and core-valence electron correlations, we computed the MUEs of CCSD(T)/aug-cc-pCVXZ (X = D, T) with and without a frozen core. The differences between them with both basis sets are around 0.4 kJ/mol, which means the influence of the core and core-valence electron correlations is not significant. Thus, it is reasonable to just consider the valence electron correlations in the calculations, and the resultant errors will be negligible.

4.3.6 Overall Performance of CCSD(T)$_{\overline{\text{F12}}}$ for Reaction Energies

The overall errors of the CCSD(T)$_{\overline{\text{F12}}}$ model chemistry for the HJO12 reaction energy set[1] are listed in Table 4.6. The errors are relative to the experimentally derived values from Ref. 1, which are obtained by subtracting the vibrational and scalar relativistic energy corrections from the experimental reaction energies. Note that we employed *valence* conventional and explicitly correlated CCSD and CCSD(T) methods; hence the core correlation contributions from Ref. 1 were included additively.

As expected, the inclusion of the F12 correction into the CCSD(T) method significantly improves the accuracy of the reaction energies. Similarly to the reaction barriers, the introduction of F12 terms reduces the basis set requirements by between 1 and 2 cardinal numbers; e.g., we found that the errors of the triple-ζ CCSD(T)$_{\overline{\text{F12}}}$ calculations are smaller than those of the quadruple-ζ CCSD(T) calculations. With the quadruple-ζ basis sets, the CCSD(T)$_{\overline{\text{F12}}}$ calculations have MUEs of about 1 kJ/mol, and MaxUEs around 3$\sim$4 kJ/mol. Remarkably,

Table 4.7: The mean unsigned errors and maximum unsigned errors (kJ/mol) of atomization energies for the HEAT test set.[a]

Method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
			cc-pCVXZ-F12			
CCSD	75.38	45.39	33.71	152.14	91.82	70.51
CCSD(2)$_{\overline{\text{F12}}}$	31.55	27.48	25.92	62.03	57.87	55.22
CCSD(T)	52.65	20.50	8.18	100.36	35.55	14.30[b]
CCSD(T)$_{\overline{\text{F12}}}$	8.82	2.62	0.95	15.88[c]	5.16[d]	2.87[d]
			aug-cc-pCVXZ			
CCSD	88.30	46.22	33.09	167.51	98.72	70.84
CCSD(2)$_{\overline{\text{F12}}}$	29.91	26.28	25.47	55.54[c]	54.28	54.20
CCSD(T)	69.85	21.81	7.60	133.14[c]	43.58	15.36[b]
CCSD(T)$_{\overline{\text{F12}}}$	11.46	2.13	0.99	29.46[c]	5.84[d]	3.14

[a]Unless noted specifically, the maximal errors were from calculations of CO_2. [b] N_2. [c] C_2H_2. [d] CN.

the aug-cc-pVXZ basis sets are found to yield smaller errors than the cc-pVXZ-F12 basis sets. This is especially pronounced for the maximum errors. In fact, the aug-cc-pVTZ basis set gives a slightly smaller MUE and MaxUE than the cc-pVQZ-F12 basis set.

4.3.7 Overall Performance of CCSD(T)$_{\overline{\text{F12}}}$ for Atomization Energies and Enthalpies of Formation

The performance of the CCSD(T)$_{\overline{\text{F12}}}$ model chemistry for thermochemical computations was further tested against the HEAT test set. We computed the atomization energies and enthalpies of formation with the *all-electron* conventional and explicitly correlated CCSD and CCSD(T) methods. The results were augmented with the (anharmonic) zero-point energies taken from the HEAT reference database; all other corrections were neglected.

The mean and maximum unsigned errors of atomization energies are listed in Table 4.7. The atomization energies computed within the HEAT model were used as the reference. The inclusion of F12 terms again reduces the errors by an equivalent of 1.5 to 2 cardinal numbers, with better improvements observed for the cc-pCVXZ-F12 series. In almost all instances, the cc-pCVXZ-F12 CCSD(T)$_{\overline{\text{F12}}}$ errors are smaller than the corresponding aug-cc-pCVXZ errors (the lone exception is the mean error of the cc-pCVTZ-F12 data).

Similar trends were observed for the enthalpies of formation (see Table 4.8). The reference values are the enthalpies of formation computed within the HEAT model using the elemental reaction approach. It can be seen that the MaxUEs come from either the C/N atoms or

Table 4.8: The mean unsigned errors and maximum unsigned errors (kJ/mol) of enthalpies of formation for the HEAT test set.[a]

Method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
			cc-pCVXZ-F12			
CCSD	29.93	21.17	17.68	83.09	52.55	41.20[b]
CCSD(2)$_{\overline{\text{F12}}}$	13.74	15.26	15.19	35.66	34.08	32.66
CCSD(T)	17.77	7.86	3.11	52.74	19.84	7.80
CCSD(T)$_{\overline{\text{F12}}}$	2.17	0.89	0.90	5.75[c]	2.47[c]	2.46[d]
			aug-cc-pCVXZ			
CCSD	34.24	23.31	18.26	98.05	57.99	41.71
CCSD(2)$_{\overline{\text{F12}}}$	11.15	14.90	15.18	29.87	32.58	32.25
CCSD(T)	24.10	9.47	3.48	73.21	26.08	8.51
CCSD(T)$_{\overline{\text{F12}}}$	2.79	1.27	1.11	11.52[e]	3.07[f]	2.71[d]

[a]The molecule, for which the error is maximal, is the C atom unless noted specifically. H_2, N_2, O_2, F_2, and CO were not included in the analysis since the elemental reaction approach, which was used to calculate the enthalpies of formation, defines these molecules as the standard states for elements H, N, O, F, and C, respectively. [b]CCH. [c]N. [d]CN. [e]C_2H_2. [f]CO_2.

molecules which contain them. In all cases, the cc-pCVXZ-F12 CCSD(T)$_{\overline{\text{F12}}}$ errors are smaller than the corresponding aug-cc-pCVXZ errors.

4.4 Conclusions

In this work, we evaluated the performance of the perturbative explicitly correlated coupled-cluster method, CCSD(T)$_{\overline{\text{F12}}}$, in combination with the two basis set families commonly used for the explicitly correlated coupled-cluster calculations, namely, cc-pVXZ-F12 and aug-cc-pVXZ. The results of the reaction barriers, reaction energies, atomization energies, and enthalpies of formation show that the inclusion of the perturbative F12 correction significantly reduce the basis set errors of the correlation energies. The effects of the F12 terms were least substantial for reaction barriers (DZ F12 errors were comparable to the TZ standard errors, i.e., an extra gain of one cardinal number) and the greatest for atomization energies and enthalpies of formation (a gain of two cardinal numbers). Overall, the performance of the two families was similar at the triple- and quadruple-ζ levels, with some differences observed at the double-ζ level.

We conclude that the aug-cc-pVDZ CCSD(T)$_{\overline{\text{F12}}}$ model chemistry is an excellent choice for computing electronic reaction barriers. With the aug-cc-pVDZ basis, the mean unsigned

error is only 1.72 kJ/mol (that surpasses all model chemistries considered by Zheng et al.[7]) and the maximum error is only 4.44 kJ/mol. The computational cost of the aug-cc-pVDZ CCSD(T)$_{\overline{\text{F12}}}$ computation is only a small factor greater than that of the corresponding standard aug-cc-pVDZ CCSD(T) energy computation and dramatically lower than that of the aug-cc-pVTZ CCSD(T) energy computation (hence, the F12 approach is clearly superior to the basis set extrapolation). While at least an order of magnitude more expensive than routine density functional model chemistries, the aug-cc-pVDZ CCSD(T)$_{\overline{\text{F12}}}$ predicts the barriers in the DBH24/08 database far more accurately, in the range of chemical accuracy.

For the electronic reaction energy calculations, the aug-cc-pVTZ CCSD(T)$_{\overline{\text{F12}}}$ model is shown to be the best choice. The mean unsigned error of the CCSD(T)$_{\overline{\text{F12}}}$ calculations with the aug-cc-pVTZ basis set is 1.5 kJ/mol, and its maximum unsigned error is 3.6 kJ/mol, which are below the chemical-accuracy mark. While the computational cost of the aug-cc-pVTZ CCSD(T)$_{\overline{\text{F12}}}$ calculations is significantly less than that of the aug-cc-pVQZ CCSD(T) calculations, the aug-cc-pVTZ CCSD(T)$_{\overline{\text{F12}}}$ model gives a more accurate prediction compared to the aug-cc-pVQZ CCSD(T) model.

The accurate predictions of atomization energies (within the range of chemical accuracy) require the CCSD(T)$_{\overline{\text{F12}}}$ method with the triple-ζ basis sets. The aug-cc-pCVTZ basis set gives slightly smaller mean unsigned errors than the cc-pCVTZ-F12 basis set (2.13 vs 2.62 kJ/mol), but yields a slightly larger maximum unsigned error compared to the cc-pCVTZ-F12 basis set (5.84 vs 5.16 kJ/mol). On the other hand, the more expensive quadruple-ζ CCSD(T) calculations yield errors about three times larger than the errors of the triple-ζ CCSD(T)$_{\overline{\text{F12}}}$ calculations. Increasing the basis set from triple-ζ to quadruple-ζ, the errors of the CCSD(T)$_{\overline{\text{F12}}}$ calculations decrease to under the chemical-accuracy bar.

The cc-pCVDZ-F12 CCSD(T)$_{\overline{\text{F12}}}$ method is preferred for the calculations of enthalpies of formation. It has a mean unsigned error of 2.17 kJ/mol and a maximum unsigned error of 5.75 kJ/mol, which are in the range of chemical accuracy. The similar accuracy requires the CCSD(T) calculations with the cc-pCVQZ-F12 basis set, which are far more expensive. The errors of the CCSD(T)$_{\overline{\text{F12}}}$ calculations for enthalpies of formation can be further reduced to below chemical accuracy with the triple-ζ basis sets. By switching to the CCSD(T)$_{\overline{\text{F12}}}$ method, we reduce the computational cost for the calculations of enthalpies of formation and improve the accuracy of the predictions at the same time.

4.5 Supporting Information

Same-spin and different-spin intermediates V, X, C, and B are obtained from *basic* intermediates $\mathcal{V}$, $\mathcal{X}$, $\mathcal{C}$, and $\mathcal{B}$ given in terms of spin-free (not antisymmetrized) integrals as

follows:*

$$\mathcal{V}_{IJ}^{IJ} = \left(\frac{f_{12}}{r_{12}}\right)_{IJ}^{IJ} - r_{PQ}^{IJ} g_{IJ}^{PQ} - r_{A'M}^{JI} g_{JI}^{A'M} - r_{A'M}^{IJ} g_{IJ}^{A'M} \tag{4.32}$$

$$\mathcal{V}_{IJ}^{JI} = \left(\frac{f_{12}}{r_{12}}\right)_{IJ}^{JI} - r_{PQ}^{JI} g_{IJ}^{PQ} - r_{A'M}^{IJ} g_{JI}^{A'M} - r_{A'M}^{JI} g_{IJ}^{A'M} \tag{4.33}$$

$$\mathcal{V}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} = \left(\frac{f_{12}}{r_{12}}\right)_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} - r_{\bar{P}\bar{Q}}^{\bar{I}\bar{J}} g_{\bar{I}\bar{J}}^{\bar{P}\bar{Q}} - r_{A'\bar{M}}^{\bar{J}\bar{I}} g_{\bar{J}\bar{I}}^{A'\bar{M}} - r_{A'\bar{M}}^{\bar{I}\bar{J}} g_{\bar{I}\bar{J}}^{A'\bar{M}} \tag{4.34}$$

$$\mathcal{V}_{\bar{I}\bar{J}}^{\bar{J}\bar{I}} = \left(\frac{f_{12}}{r_{12}}\right)_{\bar{I}\bar{J}}^{\bar{J}\bar{I}} - r_{\bar{P}\bar{Q}}^{\bar{J}\bar{I}} g_{\bar{I}\bar{J}}^{\bar{P}\bar{Q}} - r_{A'\bar{M}}^{\bar{I}\bar{J}} g_{\bar{J}\bar{I}}^{A'\bar{M}} - r_{A'\bar{M}}^{\bar{J}\bar{I}} g_{\bar{I}\bar{J}}^{A'\bar{M}} \tag{4.35}$$

$$\mathcal{V}_{I\bar{J}}^{I\bar{J}} = \left(\frac{f_{12}}{r_{12}}\right)_{I\bar{J}}^{I\bar{J}} - r_{P\bar{Q}}^{I\bar{J}} g_{I\bar{J}}^{P\bar{Q}} - r_{A'M}^{\bar{J}I} g_{I\bar{J}}^{MA'} - r_{A'\bar{M}}^{I\bar{J}} g_{I\bar{J}}^{A'\bar{M}} \tag{4.36}$$

$$\mathcal{V}_{I\bar{J}}^{\bar{J}I} = \left(\frac{f_{12}}{r_{12}}\right)_{I\bar{J}}^{\bar{J}I} - r_{P\bar{Q}}^{\bar{J}I} g_{I\bar{J}}^{P\bar{Q}} - r_{A'M}^{I\bar{J}} g_{\bar{J}I}^{A'M} - r_{A'\bar{M}}^{\bar{J}I} g_{I\bar{J}}^{A'\bar{M}} \tag{4.37}$$

$$\mathcal{X}_{IJ}^{IJ} = \left(f_{12}^2\right)_{IJ}^{IJ} - r_{PQ}^{IJ} r_{IJ}^{PQ} - r_{A'M}^{JI} r_{JI}^{A'M} - r_{A'M}^{IJ} r_{IJ}^{A'M} \tag{4.38}$$

$$\mathcal{X}_{IJ}^{JI} = \left(f_{12}^2\right)_{IJ}^{JI} - r_{PQ}^{JI} r_{IJ}^{PQ} - r_{A'M}^{IJ} r_{JI}^{A'M} - r_{A'M}^{JI} r_{IJ}^{A'M} \tag{4.39}$$

$$\mathcal{X}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} = \left(f_{12}^2\right)_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} - r_{\bar{P}\bar{Q}}^{\bar{I}\bar{J}} r_{\bar{I}\bar{J}}^{\bar{P}\bar{Q}} - r_{A'\bar{M}}^{\bar{J}\bar{I}} r_{\bar{J}\bar{I}}^{A'\bar{M}} - r_{A'\bar{M}}^{\bar{I}\bar{J}} r_{\bar{I}\bar{J}}^{A'\bar{M}} \tag{4.40}$$

$$\mathcal{X}_{\bar{I}\bar{J}}^{\bar{J}\bar{I}} = \left(f_{12}^2\right)_{\bar{I}\bar{J}}^{\bar{J}\bar{I}} - r_{\bar{P}\bar{Q}}^{\bar{J}\bar{I}} r_{\bar{I}\bar{J}}^{\bar{P}\bar{Q}} - r_{A'\bar{M}}^{\bar{I}\bar{J}} r_{\bar{J}\bar{I}}^{A'\bar{M}} - r_{A'\bar{M}}^{\bar{J}\bar{I}} r_{\bar{I}\bar{J}}^{A'\bar{M}} \tag{4.41}$$

$$\mathcal{X}_{I\bar{J}}^{I\bar{J}} = \left(f_{12}^2\right)_{I\bar{J}}^{I\bar{J}} - r_{P\bar{Q}}^{I\bar{J}} r_{I\bar{J}}^{P\bar{Q}} - r_{A'M}^{\bar{J}I} r_{\bar{J}I}^{A'M} - r_{A'\bar{M}}^{I\bar{J}} r_{I\bar{J}}^{A'\bar{M}} \tag{4.42}$$

$$\mathcal{X}_{I\bar{J}}^{\bar{J}I} = \left(f_{12}^2\right)_{I\bar{J}}^{\bar{J}I} - r_{P\bar{Q}}^{\bar{J}I} r_{I\bar{J}}^{P\bar{Q}} - r_{A'M}^{I\bar{J}} r_{\bar{J}I}^{A'M} - r_{A'\bar{M}}^{\bar{J}I} r_{I\bar{J}}^{A'\bar{M}} \tag{4.43}$$

$$\mathcal{X}_{\bar{J}I}^{I\bar{J}} = \left(f_{12}^2\right)_{\bar{J}I}^{I\bar{J}} - r_{P\bar{Q}}^{I\bar{J}} r_{\bar{J}I}^{P\bar{Q}} - r_{A'M}^{\bar{J}I} r_{I\bar{J}}^{A'M} - r_{A'\bar{M}}^{I\bar{J}} r_{\bar{J}I}^{A'\bar{M}} = \mathcal{X}_{I\bar{J}}^{\bar{J}I} \tag{4.44}$$

$$\mathcal{X}_{\bar{J}I}^{\bar{J}I} = \left(f_{12}^2\right)_{\bar{J}I}^{\bar{J}I} - r_{P\bar{Q}}^{\bar{J}I} r_{\bar{J}I}^{P\bar{Q}} - r_{A'M}^{I\bar{J}} r_{I\bar{J}}^{A'M} - r_{A'\bar{M}}^{\bar{J}I} r_{\bar{J}I}^{A'\bar{M}} \tag{4.45}$$

$$\mathcal{C}_{IJ}^{AB} = F_{A'}^{A} r_{IJ}^{A'B} + F_{A'}^{B} r_{JI}^{A'A} \tag{4.46}$$

$$\mathcal{C}_{I\bar{J}}^{A\bar{B}} = F_{A'}^{A} r_{I\bar{J}}^{A'\bar{B}} + F_{\bar{A}'}^{\bar{B}} r_{\bar{J}I}^{A'A} \tag{4.47}$$

$$\mathcal{C}_{\bar{J}I}^{A\bar{B}} = F_{A'}^{A} r_{\bar{J}I}^{A'\bar{B}} + F_{\bar{A}'}^{\bar{B}} r_{I\bar{J}}^{A'A} \tag{4.48}$$

$$\mathcal{C}_{\bar{I}\bar{J}}^{\bar{A}\bar{B}} = F_{\bar{A}'}^{\bar{A}} r_{\bar{I}\bar{J}}^{A'\bar{B}} + F_{\bar{A}'}^{\bar{B}} r_{\bar{J}\bar{I}}^{A'\bar{A}} \tag{4.49}$$

---

*I, J denote active occupied (alpha-spin) orbitals, M, N – all occupied orbitals, A, B – all unoccupied orbitals expressible in the orbital basis set, P, Q – all orbitals expressible in the orbital basis set, A', B' – CABS orbitals, P', Q', R' – all orbitals. Bar denotes beta-spin spatial orbitals. g are the electron repulsion integrals, r are the integrals of the bare correlation factor $f(r_{12})$, $[r,[\hat{T},r]]$ are the integrals of the $[f(r_{12}),[\hat{T}_1+\hat{T}_2,f(r_{12})]]$ operator, F and K are the Fock and exchange operator integrals, and $hJ \equiv F+K$.

$$
\begin{aligned}
\mathcal{B}_{I\bar{J}}^{I\bar{J}} = & \frac{1}{2}\left([r[T,r]]\right)_{I\bar{J}}^{I\bar{J}} + \Big(\left(r^2\right)_{P'\bar{J}}^{I\bar{J}}(hJ)_I^{P'} + \left(r^2\right)_{P'I}^{\bar{J}I}(hJ)_{\bar{J}}^{P'}\Big) \\
& -\Big(r_{R'P'}^{I\bar{J}}\left(K_{Q'}^{P'} + K_{\bar{Q}'}^{\bar{P}'}\right) r_{I\bar{J}}^{R'Q'} \\
& + r_{P'\bar{M}}^{I\bar{J}} F_{Q'}^{P'} r_{I\bar{J}}^{Q'\bar{M}} + r_{P'M}^{\bar{J}I} F_{\bar{Q}'}^{\bar{P}'} r_{\bar{J}I}^{Q'M} \\
& + r_{P\bar{A}}^{I\bar{J}} F_{Q}^{P} r_{I\bar{J}}^{Q\bar{A}} + r_{\bar{P}A}^{\bar{J}I} F_{\bar{Q}}^{\bar{P}} r_{\bar{J}I}^{\bar{Q}A} \\
& - r_{A'\bar{M}}^{I\bar{J}} F_{\bar{N}}^{\bar{M}} r_{I\bar{J}}^{A'\bar{N}} - r_{A'M}^{\bar{J}I} F_{N}^{M} r_{\bar{J}I}^{A'N}\Big) \\
& -2\Big(r_{A'\bar{M}}^{I\bar{J}} F_{\bar{P}'}^{\bar{M}} r_{I\bar{J}}^{A'P'} + r_{A'M}^{\bar{J}I} F_{P'}^{M} r_{\bar{J}I}^{A'P'} \\
& + r_{P\bar{A}}^{I\bar{J}} F_{A'}^{P} r_{I\bar{J}}^{A'\bar{A}} + r_{\bar{P}A}^{\bar{J}I} F_{\bar{A}'}^{\bar{P}} r_{\bar{J}I}^{A'A}\Big)
\end{aligned}
\tag{4.50}
$$

$$
\begin{aligned}
\mathcal{B}^{\bar{J}I}_{I\bar{J}} = & \frac{1}{2}\left([r[T,r]]\right)^{\bar{J}I}_{I\bar{J}} + \left(\left(r^2\right)^{\bar{J}I}_{P'\bar{J}}(hJ)^{P'}_{I} + \left(r^2\right)^{I\bar{J}}_{P'I}(hJ)^{P'}_{\bar{J}}\right) \\
& -\Big(r^{\bar{J}I}_{R'P'}\left(K^{P'}_{Q'} + K^{\bar{P}'}_{\bar{Q}'}\right)r^{R'Q'}_{I\bar{J}} \\
& \quad + r^{\bar{J}I}_{P'\bar{M}}F^{P'}_{Q'}r^{Q'\bar{M}}_{I\bar{J}} + r^{I\bar{J}}_{P'M}F^{\bar{P}'}_{\bar{Q}'}r^{Q'M}_{\bar{J}I} \\
& \quad + r^{\bar{J}I}_{P\bar{A}}F^{P}_{Q}r^{Q\bar{A}}_{I\bar{J}} + r^{I\bar{J}}_{\bar{P}A}F^{\bar{P}}_{\bar{Q}}r^{\bar{Q}A}_{\bar{J}I} \\
& \quad - r^{\bar{J}I}_{A'\bar{M}}F^{\bar{M}}_{\bar{N}}r^{A'\bar{N}}_{I\bar{J}} - r^{I\bar{J}}_{A'M}F^{M}_{N}r^{A'N}_{\bar{J}I}\Big) \\
& -2\Big(r^{\bar{J}I}_{A'\bar{M}}F^{\bar{M}}_{\bar{P}'}r^{A'P'}_{I\bar{J}} + r^{I\bar{J}}_{A'M}F^{M}_{P'}r^{A'P'}_{\bar{J}I} \\
& \quad + r^{\bar{J}I}_{P\bar{A}}F^{P}_{A'}r^{A'\bar{A}}_{I\bar{J}} + r^{I\bar{J}}_{\bar{P}A}F^{\bar{P}}_{\bar{A}'}r^{A'A}_{\bar{J}I}\Big)
\end{aligned} \tag{4.51}
$$

$$
\begin{aligned}
\mathcal{B}^{I\bar{J}}_{\bar{J}I} = & \frac{1}{2}\left([r[T,r]]\right)^{\bar{J}I}_{I\bar{J}} + \left(\left(r^2\right)^{\bar{J}I}_{P'\bar{J}}(hJ)^{P'}_{I} + \left(r^2\right)^{I\bar{J}}_{P'I}(hJ)^{P'}_{\bar{J}}\right) \\
& -\Big(r^{I\bar{J}}_{R'P'}\left(K^{P'}_{Q'} + K^{\bar{P}'}_{\bar{Q}'}\right)r^{R'Q'}_{\bar{J}I} \\
& \quad + r^{I\bar{J}}_{P'\bar{M}}F^{P'}_{Q'}r^{Q'\bar{M}}_{\bar{J}I} + r^{\bar{J}I}_{P'M}F^{\bar{P}'}_{\bar{Q}'}r^{Q'M}_{I\bar{J}} \\
& \quad + r^{I\bar{J}}_{P\bar{A}}F^{P}_{Q}r^{Q\bar{A}}_{\bar{J}I} + r^{\bar{J}I}_{\bar{P}A}F^{\bar{P}}_{\bar{Q}}r^{\bar{Q}A}_{I\bar{J}} \\
& \quad - r^{I\bar{J}}_{A'\bar{M}}F^{\bar{M}}_{\bar{N}}r^{A'\bar{N}}_{\bar{J}I} - r^{\bar{J}I}_{A'M}F^{M}_{N}r^{A'N}_{I\bar{J}}\Big) \\
& -2\Big(r^{I\bar{J}}_{A'\bar{M}}F^{\bar{M}}_{\bar{P}'}r^{A'P'}_{\bar{J}I} + r^{\bar{J}I}_{A'M}F^{M}_{P'}r^{A'P'}_{I\bar{J}} \\
& \quad + r^{I\bar{J}}_{P\bar{A}}F^{P}_{A'}r^{A'\bar{A}}_{\bar{J}I} + r^{\bar{J}I}_{\bar{P}A}F^{\bar{P}}_{\bar{A}'}r^{A'A}_{I\bar{J}}\Big)
\end{aligned} \tag{4.52}
$$

$$
\begin{aligned}
\mathcal{B}^{\bar{J}I}_{\bar{J}I} = & \frac{1}{2}\left([r[T,r]]\right)^{\bar{J}I}_{\bar{J}I} + \left(\left(r^2\right)^{I\bar{J}}_{P'\bar{J}}(hJ)^{P'}_{I} + \left(r^2\right)^{\bar{J}I}_{P'I}(hJ)^{P'}_{\bar{J}}\right) \\
& -\Big(r^{\bar{J}I}_{R'P'}\left(K^{P'}_{Q'} + K^{\bar{P}'}_{\bar{Q}'}\right)r^{R'Q'}_{\bar{J}I} \\
& \quad + r^{\bar{J}I}_{P'\bar{M}}F^{P'}_{Q'}r^{Q'\bar{M}}_{\bar{J}I} + r^{I\bar{J}}_{P'M}F^{\bar{P}'}_{\bar{Q}'}r^{Q'M}_{I\bar{J}} \\
& \quad + r^{\bar{J}I}_{P\bar{A}}F^{P}_{Q}r^{Q\bar{A}}_{\bar{J}I} + r^{I\bar{J}}_{\bar{P}A}F^{\bar{P}}_{\bar{Q}}r^{\bar{Q}A}_{I\bar{J}} \\
& \quad - r^{\bar{J}I}_{A'\bar{M}}F^{\bar{M}}_{\bar{N}}r^{A'\bar{N}}_{\bar{J}I} - r^{I\bar{J}}_{A'M}F^{M}_{N}r^{A'N}_{I\bar{J}}\Big) \\
& -2\Big(r^{\bar{J}I}_{A'\bar{M}}F^{\bar{M}}_{\bar{P}'}r^{A'P'}_{\bar{J}I} + r^{I\bar{J}}_{A'M}F^{M}_{P'}r^{A'P'}_{I\bar{J}} \\
& \quad + r^{\bar{J}I}_{P\bar{A}}F^{P}_{A'}r^{A'\bar{A}}_{\bar{J}I} + r^{I\bar{J}}_{\bar{P}A}F^{\bar{P}}_{\bar{A}'}r^{A'A}_{I\bar{J}}\Big)
\end{aligned} \tag{4.53}
$$

Note the appearance of mixed-case alpha-beta/beta-alpha integrals in many of these expression – these appear due to the action of the spin projectors on alpha-beta pairs. Then

intermediates V for same- and opposite-spin cases are easily obtained:

$$V_{IJ}^{IJ} = \mathcal{V}_{IJ}^{IJ} - \mathcal{V}_{IJ}^{JI} \tag{4.54}$$
$$V_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} = \mathcal{V}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} - \mathcal{V}_{\bar{I}\bar{J}}^{\bar{J}\bar{I}} \tag{4.55}$$
$$V_{I\bar{J}}^{I\bar{J}} = \mathcal{V}_{I\bar{J}}^{I\bar{J}} \tag{4.56}$$
$$V_{I\bar{J}}^{\bar{J}I} = \mathcal{V}_{I\bar{J}}^{\bar{J}I} \tag{4.57}$$

Intermediates C are similar:

$$C_{IJ}^{AB} = \mathcal{C}_{IJ}^{AB} - \mathcal{C}_{JI}^{AB} \tag{4.58}$$
$$C_{\bar{A}\bar{B}}^{\bar{A}\bar{B}} = \mathcal{C}_{\bar{I}\bar{J}}^{\bar{A}\bar{B}} - \mathcal{C}_{\bar{J}\bar{I}}^{\bar{A}\bar{B}} \tag{4.59}$$
$$C_{I\bar{J}}^{A\bar{B}} = \mathcal{C}_{I\bar{J}}^{A\bar{B}} \tag{4.60}$$
$$C_{\bar{J}I}^{A\bar{B}} = \mathcal{C}_{\bar{J}I}^{A\bar{B}} \tag{4.61}$$

Computation of X and B is only slightly more complicated because both subscript and superscript (bra and ket) indices must be swapped, e.g., not only $X_{I\bar{J}}^{I\bar{J}}$ and $X_{I\bar{J}}^{\bar{J}I}$ are needed but also $X_{\bar{J}I}^{\bar{J}I}$; note that the latter differs from $X_{I\bar{J}}^{I\bar{J}}$ for open-shell systems due to the sums over the occupied orbitals. This minor issue occurs due to the spin dependence of the strong orthogonality projector used to construct the geminals. Similar conclusions can be reached for B.

Chapter 5

Anatomy of Molecular Properties Evaluated with Explicitly Correlated Electronic Wave Function Methods

5.1 Introduction

During the last few decades, rapid developments in electronic structure theory have allowed the prediction of electronic energies to reach chemical (1 kcal/mol) or even spectroscopic (1 cm^{-1}) accuracy. For example, the reaction barriers in the DBH24/08 database developed by Truhlar et al.[7] can be computed within chemical accuracy by using the coupled-cluster singles, doubles, and perturbative triples method (CCSD(T)), with the relatively large aug-cc-pCV(T+d)Z basis set. On the other hand, this demonstrates that the accurate computation of the electronic energies requires both high-level wave function models (like coupled-cluster methods) and larger basis sets. The same statements can also be made for the prediction of other molecular properties, such as the electric dipole moment. Although it is technically simple to compute, spectroscopic-accuracy prediction of the electric dipole moment demands a high-level correlation method (internally contracted multireference configuration interaction method (IC-MRCI)) combined with a large basis set.[131] For example, only with the large aug-cc-pV6Z basis set is the basis set error of the IC-MRCI dipole moment of the water molecule at its equilibrium geometry reduced below 10^{-3} a.u. In this chapter, our concern is the reduction of the basis set error in the calculations on molecular properties other than chemical energy differences and the effectiveness of using explicitly correlated methods to solve this problem.

The basis set problem of conventional wave function methods is due to the orbital product nature of their building blocks, the Slater determinants, as orbital products alone cannot efficiently describe the cusp behavior of the exact wave function at short inter-electronic dis-

tances. Explicitly correlated methods overcome this problem by including the inter-electronic distances (r_{ij}) explicitly in their wave function expansions. Today, the F12 (or R12) methods, which were pioneered by Kutzelnigg,[9] are the most practical explicitly correlated methods. Moreover, F12 methods have undergone rapid development over the past decade, and become common tools in computational chemistry for accurate prediction of ground-state energies.

Indeed, since their emergence the F12 methods have been used mainly to obtain accurate molecular energies, yet they can be extended to computations of molecular properties. Furthermore, the basis set errors of molecular properties can also be significant, thus the extension of F12 methods to computations of molecular properties can be potentially crucial for their accurate prediction. F12 calculations on molecular properties, such as dipole moments, polarizabilities, and first and second hyperpolarizabilities, can be found in a few studies,[12–14,121] where the analytic calculation of the relaxed density matrix has been implemented. At the MP2-R12 level, Kordel and co-workers[12] implemented the analytic calculation of the relaxed density matrix. Later, Klopper et al.[13] reported their implementation within the MP2-F12 framework, where a Slater-type geminal was used as the correlation factor. Efforts have also been made on extensions to the CCSD method. Neiss and Hättig[121] were the first to report the implementation of the CCSD(R12) response theory. Then, Yang and Hättig developed their CCSD(R12) and CCSD(F12) response theory,[14] where the more accurate ansatz 2 of the F12 theory was employed.

In this chapter, we assess the performance of the MP2-F12 method for computing molecular properties in detail. Our approach is based on the diagonal orbital-invariant (SP) F12 ansatz.[68] Moreover, approximation C of the MP2-F12 theory[67] was employed in the evaluation, whereas previously (the less rigorous) approximation A and approximation B have been used.[12–14,121] The method was used to compute the electric dipole and quadrupole moments of BH, HF,[132] H_2O,[131] and CO,[133] for which accurate reference data can be obtained using standard wave function methods combined with the aug-cc-pVXZ basis sets of Dunning et al.[18–20] In this work, we also included calculations with the cc-pVXZ-F12 series of F12-optimized basis sets of Peterson et al.[15–17] Our previous work has demonstrated that neither one of these two basis set series is a clear winner for prediction of different relative energies.[93] Here, we compare the performance of the two basis set families for prediction of the static electric dipole and quadrupole moments.

This chapter is organized as follows. First, in Section 5.2 we briefly review the MP2-F12 theory, and then present the formalism for the relaxed one-electron density as well as the computational details. In Section 5.3, we first analyze different contributions to the dipole and quadrupole moments, and then the basis set convergence of the Hartree-Fock (HF) and MP2 correlation contributions. Furthermore, we compare the MP2 and MP2-F12 results in this section. Finally, we summarize our findings in Section 5.4.

5.2 Computational Methods

5.2.1 The Formalism of MP2-F12 One-Electron Density

The MP2-F12 correlation energy is obtained by minimizing the Hylleraas functional,

$$H^{(2)} = 2\langle\Psi^{(1)}|\hat{V}|\Psi^{(0)}\rangle + \langle\Psi^{(1)}|\hat{H}^{(0)} - E^{(0)}|\Psi^{(1)}\rangle, \tag{5.1}$$

where the zeroth-order wave function is the HF determinant, and the first-order wave function is comprised of standard and geminal double excitations. In this work, the diagonal orbital-invariant (SP) ansatz of Ten-no[68] is adopted. Thus, the first-order wave function is written as

$$\Psi^{(1)}_{\text{MP-R12}} = \frac{1}{4}t^{ij}_{ab}\tilde{a}^{ab}_{ij}|\Psi^{(0)}\rangle + \frac{1}{4}\bar{R}^{ij}_{\alpha\beta}\tilde{a}^{\alpha\beta}_{ij}|\Psi^{(0)}\rangle, \tag{5.2}$$

where t^{ij}_{ab} are conventional MP1 amplitudes, and $\bar{R}^{ij}_{\alpha\beta}$ are the antisymmetrized integrals of the projected geminal correlation factor, $f(r_{12})$:

$$\bar{R}^{ij}_{\alpha\beta} = \langle\alpha\beta|\hat{Q}_{12}f(r_{12})|ij\rangle - \langle\alpha\beta|\hat{Q}_{12}f(r_{12})|ji\rangle. \tag{5.3}$$

Here, the projector $\hat{Q}_{12}$ is defined as

$$\hat{Q}_{12} = (1-\hat{O}_1)(1-\hat{O}_2) - \hat{V}_1\hat{V}_2, \tag{5.4}$$

where $\hat{O}$ and $\hat{V}$ are the projectors on the occupied and virtual orbitals. It ensures that the geminal functions are orthogonal to both the reference and standard double excitations.[66] The detailed explanation for the notation here can be found in Appendix A.

In the SP approach, the geminal correlation factor is determined to satisfy the singlet and triplet coalescence conditions.[68] Thus, the spin-adapted geminal correlation factor can be written as[93]

$$f(r_{12}) \equiv (C_0\hat{P}_0 + C_1\hat{P}_1)\gamma(r_{12}), \tag{5.5}$$

where $C_0 = 1/2$, $C_1 = 1/4$, and $\hat{P}_0$ and $\hat{P}_1$ are the singlet and triplet spin projectors, respectively. These spin projectors ensure that the proper coefficients are selected for the different pair functions. Here, $\gamma(r_{12})$ is the standard exponential correlation factor introduced by Ten-no.[71] In this work, it is approximated with an expansion of Gaussian geminals:[73]

$$\gamma(r_{12}) \equiv -l_c\exp(-r_{12}/l_c) \approx \sum_i^{N_g} c_i\exp(-\alpha_i r_{12}^2), \tag{5.6}$$

where l_c ($\sim 1\,a_0$) is a correlation length scale and depends empirically on the orbital basis set.

A one-electron property of a molecule can be defined as its response to an external perturbation (λ).[134] The energy of a molecule in the presence of such a perturbation is expressed as

$$E(\lambda) = E(\lambda = 0) + \lambda \frac{dE}{d\lambda}\bigg|_{\lambda=0} + \frac{1}{2}\,\lambda^2 \frac{d^2E}{d\lambda^2}\bigg|_{\lambda=0} + \cdots . \tag{5.7}$$

The corresponding one-electron property thus can be calculated as the first derivative of the electronic energy with respect to the external perturbation. The dipole moment ($\boldsymbol{\mu}$), for example, is obtained by differentiating the energy with respect to an external electric field ($\mathbf{F}$):

$$\boldsymbol{\mu} = -\frac{dE}{d\mathbf{F}}\bigg|_{\mathbf{F}=0} = D_{\kappa\lambda}\langle\kappa|\hat{\mathbf{O}}|\lambda\rangle, \tag{5.8}$$

where $\langle\kappa|\hat{\mathbf{O}}|\lambda\rangle$ are integrals over the operator $\hat{\mathbf{O}}$ associated with the property O, and $D_{\kappa\lambda}$ is the relaxed one-electron density. Without considering the orbital-relaxation effects, the (unrelaxed) one-electron density of MP2-F12 is obtained simply from $\langle\Psi_{\text{MP2-F12}}|a^{\kappa}_{\lambda}|\Psi_{\text{MP2-F12}}\rangle$. The detailed expressions for the resulting matrices are:

$$D^{\text{MP2-F12}}_{ij} = \frac{1}{2}\tilde{t}^{ik}_{ab}\tilde{t}^{ab}_{jk} + \frac{1}{2}\bar{R}^{ik}_{\alpha\beta}\bar{R}^{\alpha\beta}_{jk}, \tag{5.9}$$

$$D^{\text{MP2-F12}}_{ab} = \frac{1}{2}\tilde{t}^{ac}_{ij}\tilde{t}^{ij}_{bc} + \frac{1}{2}\bar{R}^{ac'}_{ij}\bar{R}^{ij}_{bc'}, \tag{5.10}$$

$$D^{\text{MP2-F12}}_{a'b} = \frac{1}{2}\bar{R}^{a'c}_{ij}\tilde{t}^{ij}_{bc} + \frac{1}{2}\bar{R}^{a'c'}_{ij}\bar{R}^{ij}_{bc'}, \tag{5.11}$$

$$D^{\text{MP2-F12}}_{a'b'} = \frac{1}{2}\bar{R}^{a'\gamma}_{ij}\bar{R}^{ij}_{b'\gamma}. \tag{5.12}$$

Here, $\tilde{t}^{ij}_{ab}$ are modified amplitudes defined as $\tilde{t}^{ij}_{ab} = t^{ij}_{ab} + A^{ij}_{ab}$, where the intermediate A is written as

$$A^{ij}_{ab} = \frac{C^{ij}_{ab}}{F^i_i + F^j_j - F^a_a - F^b_b} \tag{5.13}$$

with

$$C^{ij}_{ab} = \frac{1}{2}\left(F^{\alpha}_{a}\bar{R}^{ij}_{\alpha b} + F^{\alpha}_{b}\bar{R}^{ij}_{a\alpha}\right). \tag{5.14}$$

If we assume the extended Brillouin condition (EBC), C^{ij}_{ab} are zero; thus, the intermediate A vanishes. As a result, $\tilde{t}^{ij}_{ab}$ become the standard MP1 amplitudes. Hence, the EBC assumption eliminates the majority of the coupling terms that are comprised of both MP2 and F12 contributions. Then, the only remaining coupling contributions, $\bar{R}^{a'c}_{ij}t^{ij}_{bc}$, are from the first term in Eq. 5.11. Without these contributions, we obtain the coupling-free, (unrelaxed) one-electron density.

To obtain relaxed one-electron properties, the response of the Hartree-Fock orbitals to the perturbation has to be taken into account. This is achieved via solving the coupled-perturbed Hartree-Fock (CPHF), or Z-vector equations,

$$\bar{\kappa}_{bn}\left(\delta_{ab}\delta_{mn}(\varepsilon_b - \varepsilon_n) + A_{bnam}\right) = X_{am}, \tag{5.15}$$

where $\bar{\kappa}_{bn}$ are Lagrange multipliers, and $A_{bnam} = 4\langle ba|nm\rangle - \langle ba|mn\rangle - \langle bn|am\rangle$ for closed-shell systems. The right-hand side of the CPHF equations, X_{am}, is obtained by differentiating the MP2-F12 Hylleraas functional with respect to the orbital rotation parameters. In our work, the Hylleraas functional with the diagonal SP ansatz is expressed as

$$H^{(2)} = \frac{1}{4}\left(f_a^a + f_b^b - f_i^i - f_j^j\right)\tilde{t}_{ij}^{ab}\tilde{t}_{ab}^{ij} + \frac{1}{2}\left(\bar{g}_{ab}^{ij} + C_{ab}^{ij}\right)\tilde{t}_{ij}^{ab} + V_{ij}^{ij} + \frac{1}{2}B_{ij}^{ij} - \frac{1}{2}\left(F_i^i + F_j^j\right)X_{ij}^{ij}, \quad (5.16)$$

where intermediates V, B, and X familiar from the MP2-F12 formalism are defined as:

$$V_{ij}^{ij} = \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{g}_{ij}^{\alpha\beta}, \quad (5.17)$$

$$B_{ij}^{ij} = \bar{R}_{\alpha\beta}^{ij}F_{\gamma}^{\beta}\bar{R}_{ij}^{\alpha\gamma}, \quad (5.18)$$

$$X_{ij}^{ij} = \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{R}_{ij}^{\alpha\beta}. \quad (5.19)$$

By differentiating $H^{(2)}$ in the above expression, we obtained two types of components for X_{am}: contributions only involving active (non-frozen) occupied orbitals (X_{ai}) and contributions from all occupied orbitals (X_{am}^{*}),

$$X_{ai} = 2\left(-\frac{1}{2}\left(\bar{g}_{bc}^{ak} + C_{bc}^{ak}\right)\tilde{t}_{il}^{bc} - V_{ik}^{ak} - V_{ak}^{ik} + \left(F_i^i + F_o^o\right)X_{ik}^{ak} - B_{ik}^{ak}\right), \quad (5.20)$$

$$\begin{aligned} X_{am}^{*} = 2\Bigg(&\bar{g}_{mk}^{al}D_l^k - \bar{g}_{m\alpha}^{a\beta}D_{\beta}^{\alpha} + \frac{1}{2}\left(\bar{g}_{mb}^{kl} + F_m^{a'}\bar{R}_{a'b}^{kl} + F_b^{a'}\bar{R}_{ma'}^{kl}\right)\tilde{t}_{kl}^{ab} + \frac{1}{2}\bar{R}_{mc'}^{kl}\bar{g}_{kl}^{ac'} + \frac{1}{2}\bar{R}_{ac'}^{kl}\bar{g}_{kl}^{mc'} \\ &- \left(F_k^k + F_l^l\right)\frac{1}{2}\bar{R}_{mb'}^{kl}\bar{R}_{kl}^{ab'} + \frac{1}{2}\left(\bar{R}_{a'b'}^{kl}F_m^{a'} + \bar{R}_{ma'}^{kl}F_{a'}^{b'}\right)\bar{R}_{kl}^{ab'} + \frac{1}{2}\bar{R}_{mc'}^{kl}F_{\beta}^{a}\bar{R}_{kl}^{\beta c'}\Bigg). \end{aligned} \quad (5.21)$$

Similar to the (unrelaxed) one-electron density, the coupling-free X_{am} is obtained by neglecting terms that contain contributions from both geminal integrals and modified amplitudes ($\tilde{t}_{ij}^{ab}$).

In our evaluation, the infinite-range sum in the intermediates V, C, and X was approximated with finite sums over the complementary auxiliary basis set (CABS).[66] For the intermediate B, approximation C of Kedzuch et al.[67] as well as the CABS approach[66] were used. After integrating over the spin degrees of freedom, the spin-free expressions for these intermediates were obtained. For example,

$$\frac{1}{2}\bar{R}_{Mc'}^{kl}\bar{g}_{kl}^{Ac'} = \frac{C_0 + 3C_1}{2}r_{MC'}^{KL}g_{KL}^{AC'} + \frac{C_0 - 3C_1}{2}r_{C'M}^{KL}g_{KL}^{AC'}, \quad (5.22)$$

$$\frac{1}{2}\bar{R}_{a'c'}^{kl}F_M^{a'}\bar{R}_{kl}^{Ac'} = \frac{C_0^2 + 3C_1^2}{2}r_{A'C'}^{KL}F_M^{A'}r_{KL}^{AC'} + \frac{C_0^2 - 3C_1^2}{2}r_{C'A'}^{KL}F_M^{A'}r_{KL}^{AC'}, \quad (5.23)$$

where the r integrals are the two-electron integrals of the bare correlation factor, $\gamma(r_{12})$. The procedure for integrating over the spin coordinates is the same as the one described in Section 4.2.1, except we only considered the closed-shell reference state in this work.

In addition, we included a correction to the Hartree-Fock wave function that is obtained from single excitations into the CABS with perturbation theory (called as "CABS singles").[80] As a result, the CABS singles correction contributes to the (unrelaxed) one-electron density:

$$D_{mn}^{\mathrm{CABS}} = -t_{\alpha}^{m} t_{n}^{\alpha}, \tag{5.24}$$

$$D_{\alpha\beta}^{\mathrm{CABS}} = -t_{m}^{\alpha} t_{\beta}^{m}, \tag{5.25}$$

$$D_{m\alpha}^{\mathrm{CABS}} = -t_{\alpha}^{m}, \tag{5.26}$$

where t_i^{α} are CABS singles amplitudes. Following the same procedure described previously, the CABS singles contribution to the right-hand side of the CPHF equations can be obtained:

$$\begin{aligned} X_{am}^{\mathrm{CABS}} = 2 \big(& -F_{\alpha}^{a} t_{m}^{\alpha} + F_{m}^{n} t_{n}^{a} - \bar{g}_{m\alpha}^{an} t_{n}^{\alpha} - \bar{g}_{mn}^{a\alpha} t_{\alpha}^{n} + \bar{g}_{mn_2}^{an_1} D_{n_1 n_2}^{\mathrm{CABS}} + F_{m}^{\alpha} D_{a\alpha}^{\mathrm{CABS}} \\ & - \bar{g}_{m\beta}^{a\alpha} D_{\alpha\beta}^{\mathrm{CABS}} \big), \end{aligned} \tag{5.27}$$

with which we can solve the the CPHF equations for the CABS singles correction and obtain its orbital response contribution.

5.2.2 Computational Details

The dipole operator can be expressed as

$$\mu_{\alpha} = \sum_{A} Q_A r_{A\alpha} - \sum_{i} r_{i\alpha}, \tag{5.28}$$

where α represents a Cartesian coordinate (x, y, or z), Q is the nuclear charge, r represents a position vector, and indices A and i label the nuclei and electrons, respectively. For the electric quadrupole moment, we used the traceless moment operator defined by Buckingham:[135]

$$\Theta_{\alpha\beta} = \frac{1}{2} \sum_{A} Z_A \left(3 r_{A\alpha} r_{A\beta} - r_A^2 \delta_{\alpha\beta} \right) - \frac{1}{2} \sum_{i} \left(3 r_{i\alpha} r_{i\beta} - r_i^2 \delta_{\alpha\beta} \right). \tag{5.29}$$

where β also represents a Cartesian coordinate. For the four molecules studied, there is only one non-zero component of the dipole moment, μ_z, whereas all the diagonal components of the traceless quadrupole moment, Θ_{xx}, Θ_{yy}, and Θ_{zz}, are non-zero. Moreover, in the linear molecules the diagonal components of the quadrupole moment have the following relationship: $\Theta_{xx} = \Theta_{yy} = -1/2\Theta_{zz}$. Hence, we are only concerned with μ_z and Θ_{zz} for BH, CO, and HF. For the nonlinear H_2O molecule, Θ_{xx} and Θ_{yy} are also studied in addition to μ_z and Θ_{zz}.

The molecular geometries of BH, HF, H_2O, and CO were obtained from the literature.[131–133] We employed two orbital basis set series: the standard aug-cc-pVXZ family (X = D, T, Q, 5) designed by Dunning et al.[18–20] for computations with standard correlated methods, and the cc-pVXZ-F12 family (X = D, T, Q) designed by Peterson and co-workers[15] specifically for F12 methods. The corresponding complementary auxiliary basis sets (CABS) were

constructed by using the CABS+ approach[66] from the aug-cc-pVXZ/OptRI[84] and cc-pVXZ-F12/OptRI basis sets.[16] In our F12 calculations, the recommended correlation length scales[15] were used, and the correlation factors were expanded in terms of six Gaussian geminals.[3] In addition, the evaluation of all post-Hartree-Fock quantities utilized the robust density fitting.[118] The aug-cc-pV(X+1)Z-RI and cc-pV(X+1)Z-RI basis sets were used for the density fitting in the aug-cc-pVXZ and cc-pVXZ-F12 calculations, respectively, where X is the cardinal number of the corresponding orbital basis. All calculations were performed with the developmental version of the Massively Parallel Quantum Chemistry (MPQC) package[101] (freely available under the GPL license at http://www.mpqc.org/).

5.3 Results and Discussions

5.3.1 Direct and Orbital Response Contributions to Dipole and Quadrupole Moments

We first investigated the direct (unrelaxed) one-electron density (D^{ur}) and orbital response (D^{or}) contributions to the dipole and quadrupole moment components, μ_z and Θ_{zz}, within the CABS singles ($E_{(2)_{\mathrm{S}}}$), MP2 correlation, and F12 correlation corrections. In addition, we also studied the effects of the coupling terms in the F12 correction. In these calculations, the aug-cc-pVXZ basis sets (X = D, T, Q) were used, and the corresponding results are presented in Figure 5.1 and 5.2.

We found that the CABS singles correction to the dipole moment is only significant at the double-ζ level for BH, HF, and H_2O, whereas at the triple- and quadruple-ζ levels it is very small. For CO, even the double-ζ value of the CABS singles correction is small (-0.00133 a.u.). Nevertheless, for all four molecules the direct density contribution is the major component in the double-ζ CABS singles correction, whereas the orbital response contribution ranges from -0.002 to -0.001 a.u. On the contrary, the orbital response effects are much more important in the MP2 correlation correction. In fact, the direct and orbital response contributions are both significant in the MP2 dipole moment correction. Lastly, we found the direct density contributions from the F12 corrections are negligible at all basis set levels. Even at the double-ζ level, their absolute values are less than 0.001 a.u. Hence, the F12 contributions result mostly from the orbital response effects. The inclusion of the coupling appears to have a small but not always negligible effect on the F12 calculations. For CO, for example, the coupling contributes 0.00367 a.u. to the total F12 correction (-0.0076 a.u.).

Similar to the dipole calculations, we found that the orbital response contribution is a relatively small portion of the CABS singles correction to the quadrupole moment for the linear molecules (see Figure 5.2(a), (b), and (d)). For HF and CO, we found the double-ζ CABS singles correction appears to be similar to or larger compared to the MP2 correlation correc-

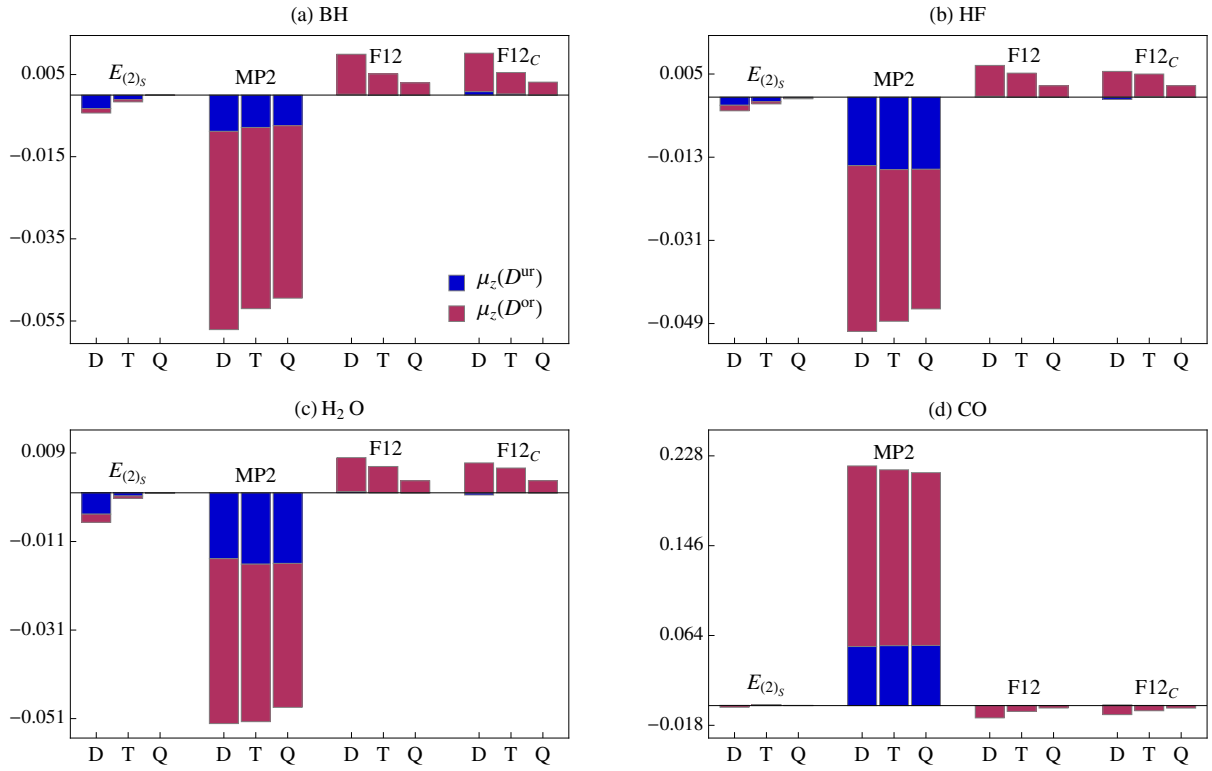

Figure 5.1: The (unrelaxed) one-electron density (D^{ur}) and orbital response (D^{or}) contributions (a.u.) from the CABS singles and correlation corrections to μ_z in the test molecules. $\mathrm{F12_C}$ and F12 refer to the F12 correlation corrections with and without the coupling from the MP2 and F12 corrections. The aug-cc-pVXZ basis sets (X = D, T, Q) were used in these calculations.

tion. This suggests that the double-ζ CABS singles correction may be more important for quadrupole moment calculations than dipole moment calculations. In the MP2 correlation correction to Θ_{zz}, the orbital response contribution increases with the size of the basis set, and it becomes much larger than the direct density contribution at the triple- and quadruple-ζ levels for BH and CO. For HF, orbital response effects at the double-ζ level are almost negligible (on the order of 0.0001 a.u.), but their contribution becomes comparable to the direct density contribution at the quadruple-ζ level. Again, for these linear molecules most of the F12 correlation contributions come from the orbital response effects ($> 90\%$ for F12 and $> 70\%$ for $\mathrm{F12_C}$).

For H_2O, observations, similar to Θ_{zz} of the HF molecule, can be made for the z^2 component of the quadrupole moment (Figure 5.2(c)): (1) the double-ζ CABS singles correction is significant and much larger than the corresponding MP2 correlation correction (-0.01027 vs. -0.00167 a.u.), and the orbital response effects are negligible for the CABS singles correction; (2) the orbital response effects on the MP2 correlation correction are still very small at the

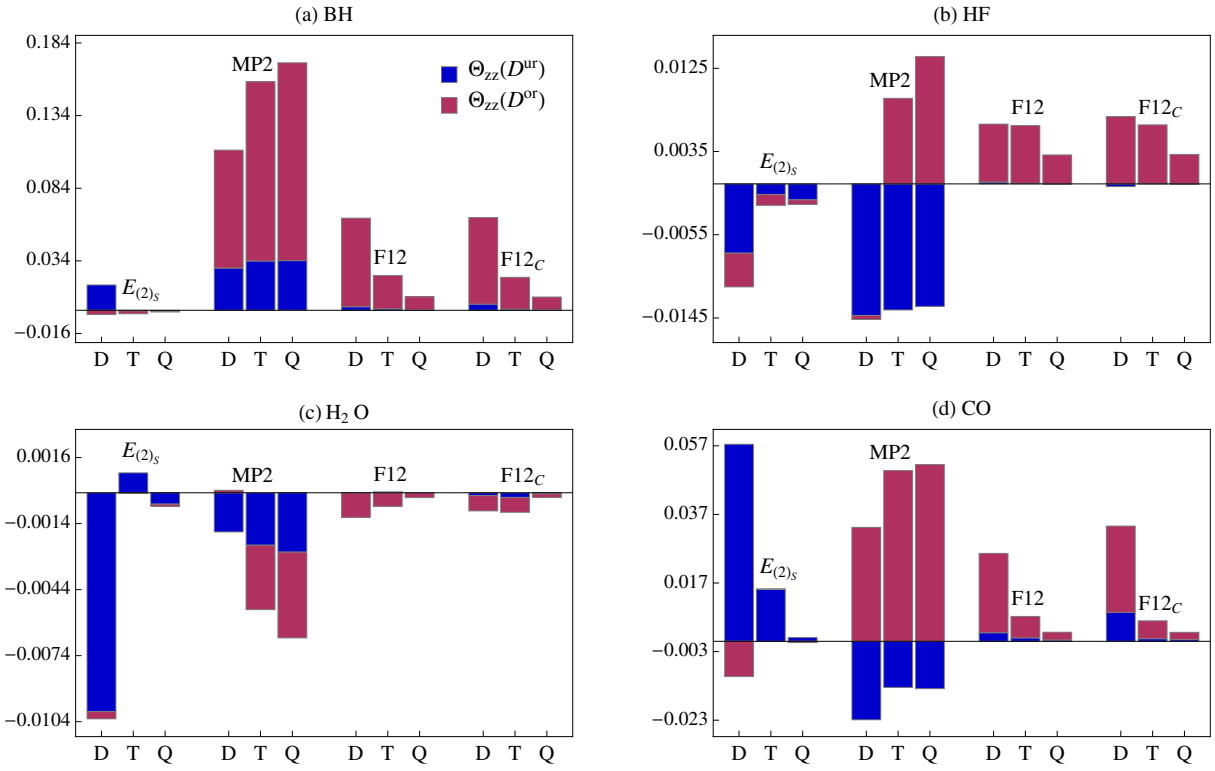

Figure 5.2: The (unrelaxed) one-electron density (D^{ur}) and orbital response (D^{or}) contributions (a.u.) from the CABS singles and correlation corrections to Θ_{zz} in the test molecules. $F12_C$ and F12 refer to the F12 correlation corrections with and without the coupling from the MP2 and F12 corrections. The aug-cc-pVXZ basis sets (X = D, T, Q) were used in these calculations.

double-ζ level (0.00011 a.u.), but become comparable to the the direct density contribution at higher levels; (3) in the F12 contributions, the orbital response effects dominate. On the other hand, the behaviors of the x^2 and y^2 components resemble those of Θ_{zz} for BH. The exceptions here are their F12 correlation corrections (see the Supporting Information in this chapter), which do not decrease monotonically with respect to the basis set size. Nevertheless, the majority of the F12 contributions for Θ_{xx} and Θ_{yy} still results from the orbital response effects.

5.3.2 Basis Set Convergence of HF and MP2 Correlation Contributions to the z Component of the Dipole Moment

To investigate the basis set convergence of various contributions to the dipole moment component, μ_z, we used two basis set families: aug-cc-pVXZ and cc-pVXZ-F12. The HF/aug-

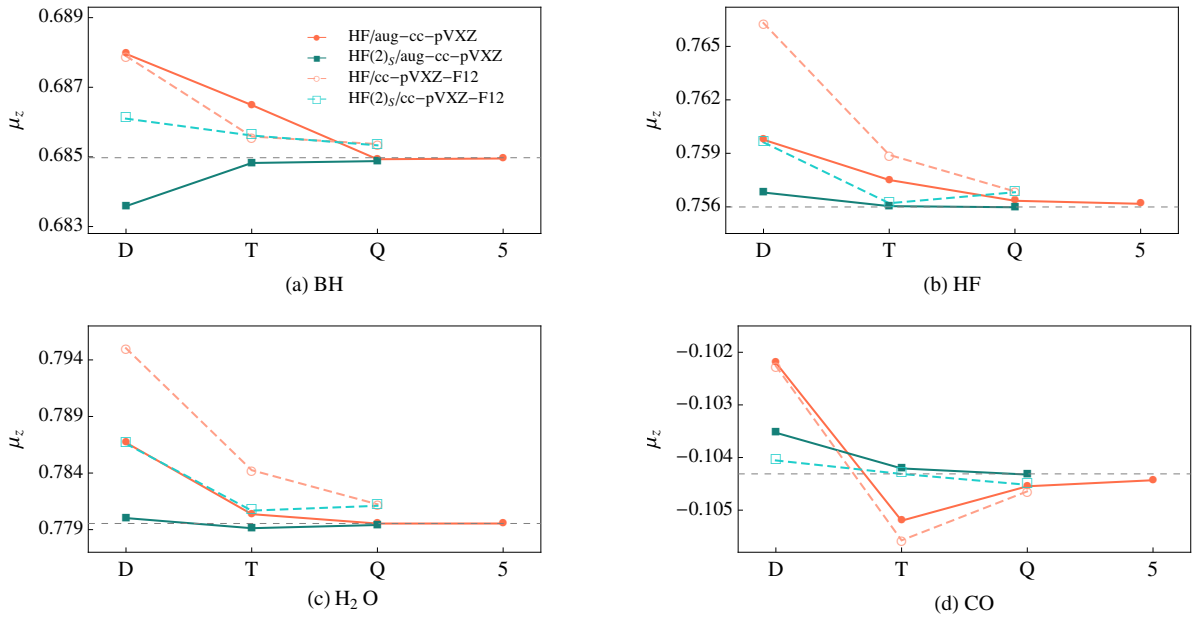

Figure 5.3: Basis set convergence of the HF and $\mathrm{HF(2)_S}$ contributions (a.u.) to μ_z for the test molecules. $\mathrm{HF(2)_S}$ refers to the Hartree-Fock contribution with the inclusion of the CABS singles correction.

cc-pV6Z values were used as the Hartree-Fock basis set limits, and the basis set limits for the MP2 correlation corrections were estimated using the standard X^{-3} extrapolation formula[33] from the aug-cc-pVQZ and aug-cc-pV5Z values.

The basis set convergence of the HF and $\mathrm{HF(2)_S}$ (HF with the CABS singles correction) contributions to the dipole moments of the four molecules are shown in Figure 5.3. For BH, HF, and CO, the CABS singles correction reduces the HF basis set errors considerably at the double- and triple-ζ levels. With the aug-cc-pVTZ basis set, the $\mathrm{HF(2)_S}$ calculations yield values very close to the basis set limits. For H_2O, even though we saw a relatively fast convergence of the HF contribution to the basis set limit, the HF basis set errors are still large at the double-ζ level (see Figure 5.3(c)). Yet the error approaches zero with the $\mathrm{HF(2)_S}$/aug-cc-pVDZ calculation. In general, the aug-cc-pVXZ family leads to a better convergence to the basis set limit in the dipole calculations. The exceptions are the cc-pVDZ-F12 $\mathrm{HF(2)_S}$ dipole moments for BH and CO, which are closer to the basis set limits than the corresponding aug-cc-pVDZ values.

In Figure 5.4, we present the basis set convergence of various correlation contributions to the z component of the dipole moment for the four molecules, where $\mathrm{MP2\text{-}F12_C}$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. As expected, the MP2 correlation contribution from the aug-cc-pVXZ calculations converges very slowly to the basis set limit, but the F12 correction

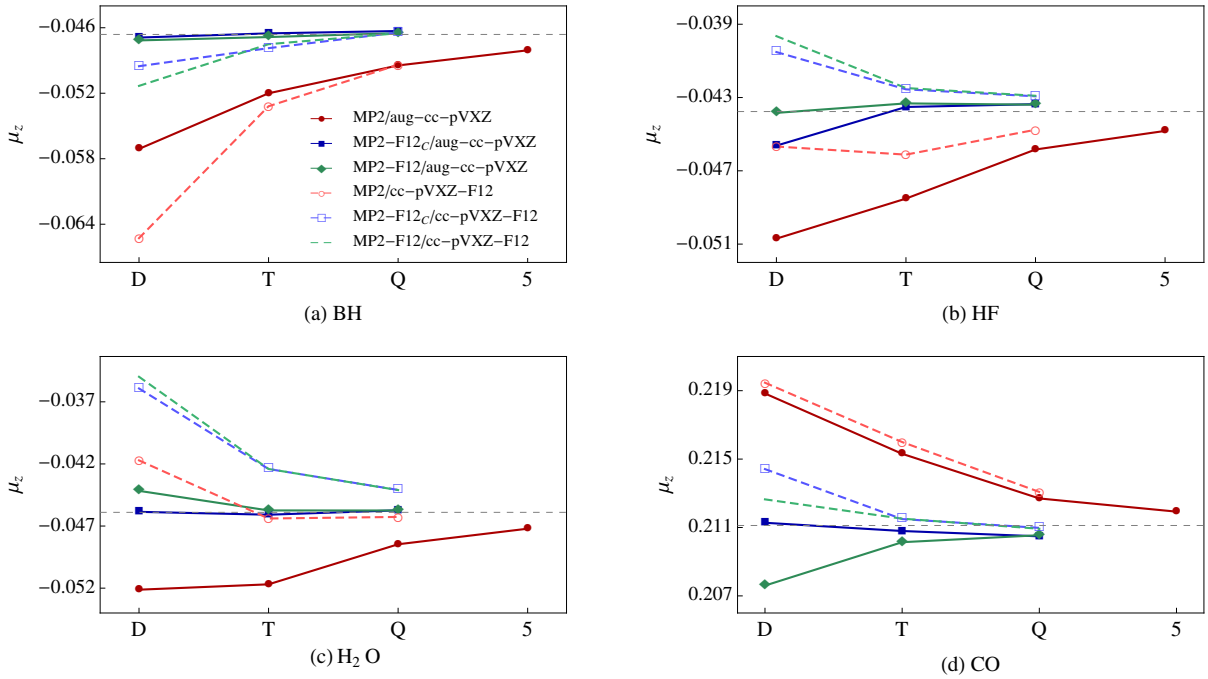

Figure 5.4: Basis set convergence of various correlation contributions (a.u.) to μ_z for the test molecules. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections.

significantly improves the basis set convergence. For BH, H_2O, and CO, the results from the MP2-F12$_C$/aug-cc-pVDZ calculations are already very close to the basis set limits, and are even better than the MP2/aug-cc-pV5Z values. Moreover, the coupling has a negligible effect on the aug-cc-pVXZ calculations of the BH dipole moment: the MP2-F12 values are almost identical to the corresponding MP2-F12$_C$ values. On the other hand, for H_2O and CO the MP2-F12 calculations yield larger errors than the MP2-F12$_C$ calculations at the double-ζ level, but the difference between the two quickly diminishes as the basis set size increases. In contrast to calculations on other molecules, the MP2-F12/aug-cc-pVXZ calculations on the HF molecule give values that converge faster to the basis set limit. Moreover, the MP2-F12/aug-cc-pVDZ value for HF is already very close to the basis set limit.

Within the MP2-F12$_C$ calculations, we found the aug-cc-pVXZ basis sets give noticeably better results than the cc-pVXZ-F12 basis sets at the double-ζ level, but their difference becomes much smaller at the triple- and quadruple-ζ levels. A similar conclusion can also be drawn for the MP2-F12 calculations of the BH, HF, and H_2O dipole moments. For CO, however, the combination of MP2-F12 with cc-pVXZ-F12 gives a better result than the MP2-F12/aug-cc-pVXZ combination: the MP2-F12/cc-pVDZ-F12 value is already very close to the basis set limit. Nevertheless, the MP2-F12$_C$ calculations with the aug-cc-pVXZ basis sets give the most consistent and accurate results, especially at the double-ζ level.

Table 5.1: The basis set errors of the MP2 correlation contributions to μ_z (a.u.) of BH, HF, and H_2O with the cc-pVXZ-F12 and modified basis sets.

Molecule	Method	cc-pVXZ-F12			(aug-)cc-pVXZ-F12[a]		
		X = D	X = T	X = Q	X = D	X = T	X = Q
BH	$MP2\text{-}F12_C$	−0.00293	−0.00128	0.00016	−0.00170	−0.00096	0.00019
	MP2-F12	−0.00473	−0.00091	0.00009	−0.00277	−0.00043	0.00017
HF	$MP2\text{-}F12_C$	0.00324	0.00121	0.00083	0.00008	0.00056	0.00062
	MP2-F12	0.00411	0.00128	0.00085	0.00128	0.00072	0.00070
H_2O	$MP2\text{-}F12_C$	0.00997	0.00349	0.00180	0.00164	0.00030	0.00040
	MP2-F12	0.01091	0.00349	0.00179	0.00331	0.00056	0.00048

[a]The cc-pVXZ-F12 basis sets were used for non-hydrogen atoms, while the aug-cc-pVXZ basis sets were used for hydrogen atoms.

Although this is a small test set, we were surprised with the overall worse performance of the cc-pVXZ-F12 series as compared to the aug-cc-pVXZ series. It is especially obvious for the dipole calculations on H_2O, which contains two hydrogen atoms. At the double-ζ level, the differences between the F12-corrected correlation values with the two basis set series are around 0.01 a.u. for H_2O, whereas the corresponding differences for the other molecules are around or less than 0.005 a.u. We know the cc-pVXZ-F12 basis sets for non-hydrogen atoms is considerably larger than the corresponding aug-cc-pVXZ basis sets, but for the hydrogen atom they are smaller when X = T, Q (cc-pVTZ-F12 has one less d function than aug-cc-pVTZ, and cc-pVQZ-F12 has one less d and f functions comparing to aug-cc-pVQZ). Hence, it is possible that the relatively poorer performance of the cc-pVXZ-F12 basis sets for dipole calculations is due to their hydrogen basis functions. To investigate the effects of hydrogen basis functions on the dipole calculations, we performed the cc-pVXZ-F12 calculations with the aug-cc-pVXZ basis sets for hydrogen atoms (labeled as (aug-)cc-pVXZ-F12 in Table 5.1). We found that the basis set errors of the calculations do become significantly smaller (see Table 5.1), and are very close to the aug-cc-pVXZ results especially at the double-ζ level (see the Supporting Information for the comparison of the three). For H_2O, the basis set errors of the correlation contributions are reduced by one order of magnitude when using the aug-cc-pVXZ basis sets for the hydrogen atoms. The results here suggest that the cc-pVXZ-F12 basis sets for hydrogen may need to be revisited as they seem to be the cause of the relatively poor results for the dipole moment.

5.3.3 Basis Set Convergence of HF and MP2 Correlation Contributions to the z^2 Component of the Quadrupole Moment

Following the same approach, we studied the basis set convergence of the HF and correlation contributions to the z^2 component of the quadrupole moment for the four test molecules. Thus, the Hartree-Fock basis set limits were also the HF/aug-cc-pV6Z values, while the basis set limits for the MP2 correlation contributions were the extrapolated values from the aug-cc-pVQZ and aug-cc-pV5Z results using the standard X^{-3} formula.[33]

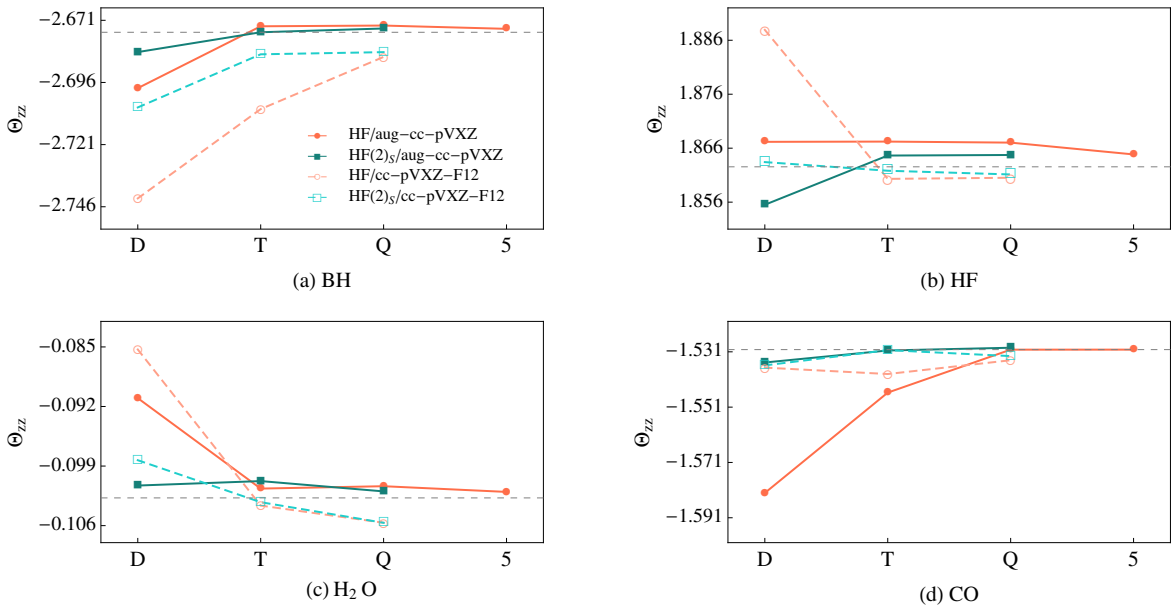

Figure 5.5: Basis set convergence of the HF and HF(2)$_S$ contributions (a.u.) to Θ_{zz} for the test molecules. HF(2)$_S$ refers to the Hartree-Fock contribution with the inclusion of the CABS singles correction.

Figure 5.5 shows the basis set convergence of the HF and HF(2)$_S$ contributions to Θ_{zz} for the four molecules. For BH and H_2O, the CABS singles correction significantly reduces the HF basis set errors in the aug-cc-pVDZ calculations, and the resultant HF(2)$_S$ values are near the basis set limits. Compared to the aug-cc-pVXZ series, the use of the cc-pVXZ-F12 basis sets leads to worse results in the HF(2)$_S$ calculations for these two molecules. However, the opposite trend is observed for the HF(2)$_S$ calculation on the HF molecule, where cc-pVDZ-F12 outperforms aug-cc-pVDZ and gives a result very close to the basis set limit. In the Θ_{zz} calculations of CO, the HF(2)$_S$ method with the two basis set families yields almost identical results, and the resultant basis set errors are already close to zero at the double-ζ level. Despite the slightly worse performance of the double-ζ calculation on the HF molecule, the use of the aug-cc-pVXZ family in the HF(2)$_S$ calculations leads to a

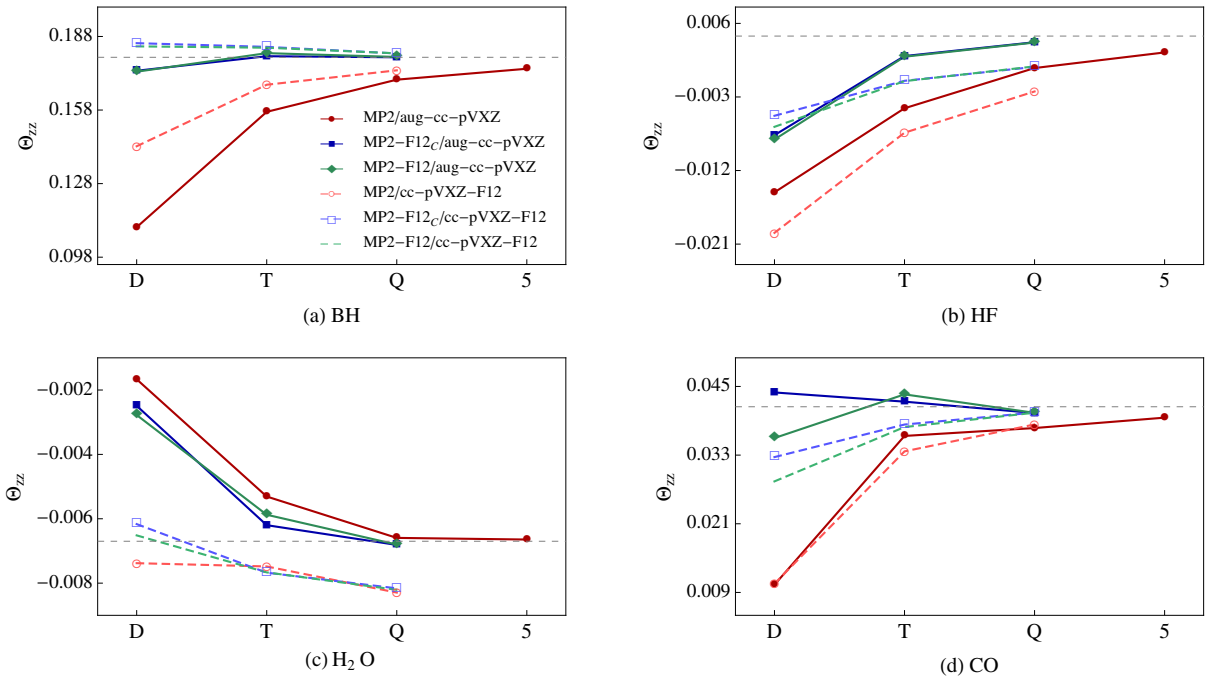

Figure 5.6: Basis set convergence of various correlation contributions (a.u.) to Θ_{zz} for the test molecules. MP2-F12$_C$ and MP2-F12 refer to the MP2 correlation contributions with and without the coupling from the MP2 and F12 corrections.

more consistent behavior of convergence compared to that of the cc-pVXZ-F12 family; this is seemingly related to the inconsistent behavior of the underlying HF/cc-pVXZ-F12 results.

The basis set convergence of the MP2 and F12-corrected correlation contributions to the z^2 component of the quadrupole moment for the test molecules are presented in Figure 5.6. For BH and CO, the basis set convergence of various correlation contributions to Θ_{zz} resembles that of the corresponding μ_z: (1) the basis set convergence of the MP2 correlation contributions is slow, but the F12 correction significantly accelerates the convergence; (2) the two basis set families generate very different results at the double-ζ level (the discrepancy is on the order of 0.01 a.u.), but the difference diminishes as we increase the basis set size. Nevertheless, the aug-cc-pVTZ basis set is needed in the MP2-F12$_C$ calculations to obtain values near the basis set limits. Similarly, the aug-cc-pVTZ basis set is also required for the MP2-F12$_C$ or MP2-F12 calculations on HF and H_2O to approach the basis set limits (within 0.001 a.u.). Although the cc-pVXZ-F12 series gives smaller errors than the aug-cc-pVXZ series in the double-ζ F12 calculations on these two molecules, they give less desirable results in the triple- and quadruple-ζ calculations. For example, the cc-pVDZ-F12 values of the F12-corrected correlation contributions to Θ_{zz} for H_2O are very close to the basis set limit, whereas the corresponding TZ and QZ values are $\sim$ 0.001 a.u. lower than the basis set limit. This is mostly due to the poor convergence of the underlying MP2/cc-pVXZ-F12

results. In addition, the MP2-F12$_C$ and MP2-F12 calculations of Θ_{zz} seem to give similar results for all four molecules, and the only obvious discrepancy is between the double-ζ calculations on CO using these two methods.

Similar to the dipole calculations, we also performed the cc-pVXZ-F12 calculations of Θ_{zz} with the aug-cc-pVXZ basis sets for hydrogen atoms. For the linear molecules (BH and HF), we found that the basis set errors of the correlation contributions become smaller ($0.001 \sim 0.003$ a.u.) and close to the aug-cc-pVXZ results (see the figures in the Supporting Information). For H_2O, the convergence of the correlation corrections to Θ_{xx} and Θ_{yy} is also much improved. However, the correlation corrections to Θ_{zz} still converge to a value much lower than the basis set limit. Since there are three Cartesian contributions, $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle z^2 \rangle$ (also called second moment of charge), to the electronic parts of Θ_{xx}, Θ_{yy}, and Θ_{zz} (see Eq. 5.29), their different convergence behaviors may be due to the inconsistent convergence of those Cartesian contributions.

Table 5.2: The MP2 correlation contributions to the components of the H_2O quadrupole moment ($\langle \alpha^2 \rangle$ in a.u.) with different basis sets.

$\langle \alpha^2 \rangle$	X	aug-cc-pVXZ		cc-pVXZ-F12		(aug-)cc-pVXZ-F12[a]	
		MP2	MP2-F12$_C$	MP2	MP2-F12$_C$	MP2	MP2-F12$_C$
$\langle x^2 \rangle$	D	−0.24053	−0.17447	−0.21292	−0.17820	−0.21775	−0.17952
	T	−0.21199	−0.18562	−0.20114	−0.17940	−0.20277	−0.18313
	Q	−0.19849	−0.18654	−0.19281	−0.18216	−0.19471	−0.18482
	X^{-3} [b]	−0.18724					
$\langle y^2 \rangle$	D	−0.19306	−0.12443	−0.17477	−0.12263	−0.17542	−0.12420
	T	−0.15300	−0.12082	−0.14940	−0.12201	−0.14671	−0.11996
	Q	−0.13499	−0.12013	−0.13523	−0.12122	−0.13297	−0.11935
	X^{-3}	−0.11963					
$\langle z^2 \rangle$	D	−0.21846	−0.15194	−0.20122	−0.15658	−0.20269	−0.15569
	T	−0.18780	−0.15942	−0.18275	−0.15838	−0.18401	−0.16112
	Q	−0.17333	−0.16015	−0.17231	−0.15971	−0.17298	−0.16116
	X^{-3}	−0.16014					

[a]The cc-pVXZ-F12 basis sets were used for the oxygen atom, while the aug-cc-pVXZ basis sets were used for the hydrogen atoms. [b]The value was obtained using the standard X^{-3} extrapolation with the aug-cc-pVQZ and aug-cc-pV5Z values.

To clarify the issue above, we computed $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle z^2 \rangle$ of H_2O with the aug-cc-pVXZ, cc-pVXZ-F12, and (aug-)cc-pVXZ-F12 basis sets. The results are presented in Table 5.2. We found that the convergences of the MP2 values are consistent: they become smaller as the basis set size increases, and slowly approach the basis set limits from below. On

the other hand, the MP2-F12$_C$ values of $\langle x^2 \rangle$ and $\langle z^2 \rangle$ with aug-cc-pVXZ and cc-pVXZ-F12 decrease with the basis set size, and thus converge to the basis set limit from above, whereas the $\langle y^2 \rangle$ values approach the basis set limit from the opposite direction. Although the MP2-F12$_C$ values of each Cartesian contributions converge relatively fast to the basis set limits, their incoherent convergence behaviors could lead to an undesired convergence for the quadrupole moments. This inconsistency of convergence among the $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle z^2 \rangle$ values is especially obvious for the results from the (aug-)cc-pVXZ-F12 calculations. Despite the fact that the basis set convergence of the MP2-F12$_C$/(aug-)cc-pVXZ-F12 values closely follows those of the corresponding aug-cc-pVXZ results, it does not always converge to the right value. This is probably more obvious from the plots in the Supporting Information (the MP2-F12 results are also included). In addition, we noticed that the $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle z^2 \rangle$ values from the MP2-F12$_C$/cc-pVXZ-F12 calculations are actually above the basis set limits at all basis set levels, even though the corresponding Θ_{zz} values converge to a value below the basis set limit. Since H_2O is a nonlinear molecule, the results here indicate that the calculations of the quadrupole moments for nonlinear molecules may be more challenging, yet more tests need to be done.

5.3.4 Comparison of Dipole and Quadrupole Moment Components, μ_z and Θ_{zz}, from MP2 and F12 Calculations

In Table 5.3 and 5.4, we list the deviations of the double-ζ μ_z values and triple-ζ Θ_{zz} values from the F12 calculations with respect to the MP2 basis set limits, which are the sums of the HF and MP2 correlation basis set limits. For the calculations of μ_z, the MP2-F12$_C$/aug-cc-pVDZ method with the CABS singles correction yields errors around 0.001 a.u. On the other hand, the errors from the MP2-F12$_C$/cc-pVDZ-F12 and MP2-F12/aug-cc-pVXZ (or cc-pVDZ-F12) calculations can be much larger for one or two molecules. In the calculations of Θ_{zz}, the errors for BH and CO can be reduced to below 0.001 a.u. with the MP2-F12$_C$/aug-cc-pVTZ model. For HF and H_2O, without the CABS singles correction the errors of MP2-F12$_C$/aug-cc-pVTZ calculations are significantly reduced, and approach 0.001 a.u. Compared with MP2-F12$_C$, MP2-F12 with aug-cc-pVTZ generates slightly larger errors (except for BH). For the Θ_{zz} calculations with both the F12 models, the cc-pVXZ-F12 basis sets give much larger errors than the aug-cc-pVXZ basis sets in most cases (H_2O is the only exception). The results for the x^2 and y^2 components of the H_2O quadrupole moment differ slightly from those of the z^2 component, but they do not provide additional insights (see the Supporting Information). Thus, for the dipole and quadrupole calculations we conclude that the MP2-F12$_C$ method combined with the aug-cc-pVXZ basis sets is the most reliable model, and the CABS singles correction is needed at the double-ζ level but not a necessity at the triple-ζ level.

Lastly, we investigated how the quality of the CABS influences the caluations of dipole and quadrupole moments. Hence, we performed the aug-cc-pVXZ (X = D, T) calculations using

Table 5.3: The errors of the double-ζ F12 calculations for μ_z (a.u.) with respect to the estimated basis set limits.[a]

Method	Basis	BH	HF	H_2O	CO
MP2-F12$_C$[b]	aug-cc-pVDZ	−0.00168	−0.00121	0.00053	0.00107
	cc-pVDZ-F12	−0.00179[c]	0.00672	0.01703	0.00368
MP2-F12[b]	aug-cc-pVDZ	−0.00194	0.00058	0.00222	−0.00260[d]
	cc-pVDZ-F12	−0.00359[e]	0.00759	0.017967	0.00191
			Basis set limit		
MP2	—	0.63836	0.71238	0.73363	0.10667

[a]The basis set limits were estimated as the sum of the HF/aug-cc-pV6Z value and the extrapolated MP2 correlation contribution from the aug-cc-pVQZ and aug-cc-pV5Z values. [b]The CABS singles correction was included in the calculations. [c]The corresponding value without the CABS singles correction is much smaller (0.00001 a.u.). [d]The value without the CABS singles correction is -0.00127 a.u. [e]The value without the CABS singles correction is -0.00179 a.u.

Table 5.4: The errors of the triple-ζ F12 calculations for Θ_{zz} (a.u.) with respect to the estimated basis set limits.[a]

Method	Basis	BH	HF	H_2O	CO
MP2-F12$_C$[b]	aug-cc-pVTZ	−0.00085	−0.00207[c]	0.00222[e]	0.00041
	cc-pVTZ-F12	−0.00591	−0.00842	−0.00174	−0.00345
MP2-F12[b]	aug-cc-pVTZ	0.00040	−0.00215[d]	0.00254[f]	0.00168
	cc-pVTZ-F12	−0.00625	−0.00846	−0.00173	−0.00390
			Basis set limit		
MP2	—	−2.49501	1.73671	−0.10918	−1.48859

[a]The basis set limits were estimated as the sum of the HF/aug-cc-pV6Z value and the extrapolated MP2 correlation contribution from the aug-cc-pVQZ and aug-cc-pV5Z values. [b]The CABS singles correction was included in the calculations. [c]The corresponding value without the CABS singles correction is much smaller (0.00026 a.u.). [d]The value without the CABS singles correction is 0.00018 a.u. [e]The corresponding value without the CABS singles correction is much smaller (0.00134 a.u.). [f]The value without the CABS singles correction is 0.00166 a.u.

Table 5.5: The variations of the CABS singles and F12 corrections to μ_z (a.u.) from the aug-cc-pVXZ calculations using different CABS with respect to the values using aug-cc-pVXZ-CABS.

Molecule	X	(uc)aug-cc-pVQZ			cc-pVXZ-F12-CABS		
		$E_{(2)\mathrm{s}}$	$\mathrm{F12_C}^a$	F12	$E_{(2)\mathrm{s}}$	$\mathrm{F12_C}^a$	F12
BH	2	0.00161	0.00016	0.00004	−0.00054	0.00001	−0.00002
	3	0.00022	0.00005	0.00010	−0.00034	−0.00002	−0.00000
HF	2	−0.00001	0.00012	0.00019	−0.00164	−0.00019	−0.00015
	3	0.00021	0.00011	0.00020	−0.00157	0.00003	0.00003
H_2O	2	0.00066	0.00016	0.00020	−0.00179	−0.00005	0.00002
	3	0.00052	0.00013	0.00025	−0.00149	0.00004	0.00006
CO	2	−0.00131	0.00031	−0.00013	−0.00012	−0.00082	−0.00074
	3	−0.00026	−0.00008	−0.00019	0.00022	−0.00006	−0.00005

[a] $\mathrm{F12_C}$ refers to the F12 correlation contribution with the coupling from the MP2 and F12 corrections.

the large uncontracted aug-cc-pVQZ and cc-pVXZ-F12-CABS (the corresponding CABS for cc-pVXZ-F12) basis sets as the CABS. In Table 5.5, we list the variations of the CABS singles and F12 corrections to μ_z from the calculations using these two CABSs, where we used the values from the calculations utilizing the regular aug-cc-pVXZ-CABS basis sets (the corresponding CABS for aug-cc-pVXZ) as the reference. We found that the results from the calculations with the different CABSs are very similar. The variation of the CABS singles correction contribution is close to or smaller than 0.001 a.u. For the F12 corrections ($\mathrm{F12_C}$ and F12) to μ_z, the effect of the choice of the CABS is even smaller: the difference among results using different CABSs is on the order of 0.0001 a.u. For the calculations of Θ_{zz} (Table 5.6), the choice of the CABS has a slightly larger influence on the CABS singles correction at the double-ζ level especially for the calculations using cc-pVDZ-F12-CABS, but the variation of the CABS singles correction with the different CABSs is again reduced to or smaller than 0.001 a.u. with the triple-ζ basis set. Similar observations can be made to the F12 corrections ($\mathrm{F12_C}$ and F12) to Θ_{zz}, and their variation with the different CABSs is very small ($\leq$ 0.0001 a.u.). Thus, we conclude that the default CABS is sufficient for the dipole moment calculations and triple- or higher-ζ quadrupole calculations.

Table 5.6: The variations of the CABS singles and F12 corrections to Θ_{zz} (a.u.) from the aug-cc-pVXZ calculations using different CABS with respect to the values using aug-cc-pVXZ-CABS.

Molecule	X	(uc)aug-cc-pVQZ			cc-pVXZ-F12-CABS		
		$E_{(2)\mathrm{s}}$	$\mathrm{F12_C}^a$	F12	$E_{(2)\mathrm{s}}$	$\mathrm{F12_C}^a$	F12
BH	2	0.00594	−0.00034	−0.00025	−0.00704	0.00114	0.00106
	3	0.00106	0.00011	0.00019	0.00164	−0.00008	−0.00008
HF	2	0.00808	0.00001	−0.00018	0.01192	−0.00077	−0.00075
	3	−0.00046	0.00008	0.00000	−0.00215	0.00006	0.00009
H_2O	2	0.00230	−0.00014	−0.00010	0.00381	−0.00033	−0.00028
	3	−0.00232	−0.00001	−0.00004	−0.00281	0.00003	0.00003
CO	2	−0.00536	−0.00100	−0.00037	0.00449	0.00111	0.00126
	3	−0.00160	0.00002	−0.00015	0.00009	0.00004	0.00003

[a] $\mathrm{F12_C}$ refers to the F12 correlation contribution with the coupling from the MP2 and F12 corrections.

5.3.5 Effects of Core and Core-Valence Electron Correlations on Dipole and Quadrupole Moment Components, μ_z and Θ_{zz}

In the dipole and quadrupole calculations above, we only included the valence electron correlations. We expect the core and core-valence electron correlations to be small since the dipole and quadrupole moment operators depend on r_i linearly and quadratically, respectively. To investigate the influence of the core and core-valence correlation effects, we compared the results from the MP2/aug-cc-pCVXZ (X = D, T, Q) calculations with and without a frozen core (Table 5.7). We found that the core and core-valence correlation effects on the calculations of μ_z, which are calculated as the difference between the frozen and non-frozen core calculations, vary for the test molecules. While these effects increase with the size of the basis set for the calculations on HF and H_2O, they become smaller for the calculations on BH and oscillate for the calculations on CO. Despite this, their values are around 0.001 a.u. for the μ_z calculations, and this is comparable to the errors of the MP2-$\mathrm{F12_C}$/aug-cc-pVDZ calculations of μ_z. Therefore, it is reasonable to conclude that the influence of the core and core-valence electron correlations on the dipole moment calculations is not significant, and calculations with only valence electrons correlated are sufficient. The core and core-valence correlation effects on the calculations of Θ_{zz} appear to be more systematic as they increase slightly with respect to the basis set size, but their magnitude varies for different molecules. While these effects are very small for the Θ_{zz} calculations on HF and H_2O (on the order of 0.0001 a.u.), they are noticeably larger for the calculations on CO ($\sim$ 0.003 a.u.). For the

Θ_{zz} calculations on BH, however, we found that the core and core-valence correlation effects are significant and cannot be neglected. Hence, there is no consistent conclusion for the core and core-valence correlation effects on the calculations of Θ_{zz}, and it may be necessary to perform calculations with all electrons correlated in some cases.

Table 5.7: The core and core-valence correlation effects (a.u.) on μ_z and Θ_{zz} for the test molecules.[a]

Property	X	BH	HF	H_2O	CO
μ_z	D	−0.00200	0.00039	0.00051	−0.00064
	T	−0.00151	0.00136	0.00187	−0.00091
	Q	−0.00035	0.00153	0.00210	−0.00035
Θ_{zz}	D	0.00434	0.00012	−0.00005	0.00226
	T	0.01347	0.00037	−0.00026	0.00273
	Q	0.01967	0.00040	−0.00030	0.00303

[a]The values were computed as the difference between the MP2/aug-cc-pCVXZ results with and without a frozen core.

5.4 Conclusions

In this work, we present the formalism for computing static one-electron properties with the MP2-F12 method that is based on modern F12 approaches, such as the CABS approach and approximation C. We tested the performance of MP2-F12 in detail for accurate computation of the electric dipole and quadrupole moments for four molecules, BH, HF, H_2O, and CO, with two basis set families. We found that the contributions from the direct (unrelaxed) density and orbital response effects are very different in the CABS singles and MP2 correlation corrections to the dipole and quadrupole moments. In the CABS singles correction, we can simply neglect the orbital response effects, whereas both the direct density and orbital response contributions are important in the MP2 correlation contribution. On the other hand, orbital response effects are the dominant contribution in the F12 correlation, whereas the direct F12 density contribution is negligibly small.

We conclude that the CABS singles correction is generally necessary at the double-ζ level for the Hartree-Fock calculations of dipole and quadrupole moments. The inclusion of the F12 correction significantly reduces the basis set errors of the MP2 correlation contributions to both dipole and quadrupole moments. With the MP2-$F12_C$/aug-cc-pVDZ calculations, we can obtain dipole moments close to the basis set limits (the errors are approximately 0.001 a.u.). For quadrupole moments, the MP2-$F12_C$/aug-cc-pVTZ calculations are needed to

approach the MP2 basis set limits. While we can significantly reduce the computational cost of the F12 correction by neglecting the coupling terms from the MP2 and F12 corrections, it does slightly increase the errors in most calculations. The performance of the aug-cc-pVXZ and cc-pVXZ-F12 basis sets usually differ significantly at the double-ζ level, but the difference becomes smaller at triple- and quadruple-ζ levels. In general, the F12 calculations with the aug-cc-pVXZ series give more accurate results comparing to those with the cc-pVXZ-F12 family; the difference can be attributed to the quality of the cc-pVXZ-F12 basis sets for the hydrogen atom.

5.5 Supporting Information

5.5.1 Detailed Expressions for Direct MP2-F12 One-Electron Density and Contributions to Right-Hand Side of CPHF Equations

For closed-shell systems, the matrix elements of the direct (unrelaxed) one-electron density are given in terms of spin-free (not antisymmetrized) integrals as follows:*

$$\begin{aligned} D_{IJ}^{\text{MP2-F12}} &= 2\,T_{AB}^{IK}T_{JK}^{AB} - T_{AB}^{KI}T_{JK}^{AB} + \frac{C_0^2+3C_1^2}{2}A_{AB}^{IK}A_{JK}^{AB} + \frac{C_0^2-3C_1^2}{2}A_{AB}^{KI}A_{JK}^{AB} \\ &+ \frac{C_0+3C_1}{2}\left(T_{AB}^{IK}A_{JK}^{AB} + A_{AB}^{IK}T_{JK}^{AB}\right) + \frac{C_0-3C_1}{2}\left(T_{AB}^{KI}A_{JK}^{AB} + A_{AB}^{KI}T_{JK}^{AB}\right) \\ &+ \frac{C_0^2+3C_1^2}{2}X_{JK}^{IK} + \frac{C_0^2-3C_1^2}{2}X_{JK}^{KI}, \end{aligned} \tag{5.30}$$

$$\begin{aligned} D_{AB}^{\text{MP2-F12}} &= 2\,T_{IJ}^{AC}T_{BC}^{IJ} - T_{IJ}^{CA}T_{BC}^{IJ} + \frac{C_0^2+3C_1^2}{2}A_{IJ}^{AC}A_{BC}^{IJ} + \frac{C_0^2-3C_1^2}{2}A_{IJ}^{CA}A_{BC}^{IJ} \\ &+ \frac{C_0+3C_1}{2}\left(T_{IJ}^{AC}A_{BC}^{IJ} + A_{IJ}^{AC}T_{BC}^{IJ}\right) + \frac{C_0-3C_1}{2}\left(T_{IJ}^{CA}A_{BC}^{IJ} + A_{IJ}^{CA}T_{BC}^{IJ}\right) \\ &+ \frac{C_0^2+3C_1^2}{2}r_{IJ}^{AC}r_{BC}^{IJ} + \frac{C_0^2-3C_1^2}{2}r_{IJ}^{CA}r_{BC}^{IJ}, \end{aligned} \tag{5.31}$$

$$\begin{aligned} D_{A'B}^{\text{MP2-F12}} &= r_{A'C}^{IJ}\left(\frac{C_0+3C_1}{2}T_{IJ}^{BC} + \frac{C_0-3C_1}{2}T_{IJ}^{CB} + \frac{C_0^2+3C_1^2}{2}A_{IJ}^{BC} + \frac{C_0^2-3C_1^2}{2}A_{IJ}^{CB}\right) \\ &+ \frac{C_0^2+3C_1^2}{2}r_{A'C}^{IJ}r_{IJ}^{BC} + \frac{C_0^2-3C_1^2}{2}r_{A'C}^{IJ}r_{IJ}^{CB}, \end{aligned} \tag{5.32}$$

$$D_{A'B'}^{\text{MP2-F12}} = \frac{C_0^2+3C_1^2}{2}r_{A'C}^{IJ}r_{IJ}^{B'C} + \frac{C_0^2-3C_1^2}{2}r_{A'C}^{IJ}r_{IJ}^{CB'}. \tag{5.33}$$

Intermediates in X_{AM} in terms of spin-free (not antisymmetrized) integrals are as follows:

$$\bar{g}_{Mk}^{Al}D_{kl}^{\text{MP2-F12}} = \left(2g_{MK}^{AL} - g_{KM}^{AL}\right)D_{KL}^{\text{MP2-F12}}, \tag{5.34}$$

$$\begin{aligned} \bar{g}_{M\beta}^{A\alpha}D_{\alpha\beta}^{\text{MP2-F12}} &= \bar{g}_{Mb}^{Ab}D_{bc}^{\text{MP2-F12}} + \bar{g}_{Mc}^{Ab'}D_{b'c}^{\text{MP2-F12}} + \bar{g}_{Mc'}^{Ab}D_{c'b}^{\text{MP2-F12}} + \bar{g}_{Mc'}^{Ab'}D_{c'b'}^{\text{MP2-F12}} \\ &= \left(2g_{MC}^{AB} - g_{KC}^{AB}\right)D_{BC}^{\text{MP2-F12}} + \left(2g_{MC}^{AB'} - g_{KC}^{AB'}\right)D_{B'C}^{\text{MP2-F12}} \\ &+ \left(2g_{MC'}^{AB} - g_{KC'}^{AB}\right)D_{C'B}^{\text{MP2-F12}} + \left(2g_{MC'}^{AB'} - g_{KC'}^{AB'}\right)D_{B'C'}^{\text{MP2-F12}}, \end{aligned} \tag{5.35}$$

*I, J denote active occupied (alpha-spin) orbitals, M, N – all occupied orbitals, A, B – all unoccupied orbitals expressible in the orbital basis set, P, Q – all orbitals expressible in the orbital basis set, A', B' – CABS orbitals, P', Q', R' – all orbitals. Bar denotes beta-spin spatial orbitals. f are the Fock integrals, g are the electron repulsion integrals, and r are the integrals of the bare correlation factor, $\gamma(r_{12})$. T_{AB}^{IJ} are the MP1 amplitudes, and $A_{AB}^{IJ} = \left(f_A^{A'}r_{A'B}^{IJ} + f_B^{A'}r_{AA'}^{IJ}\right)/\left(f_I^I + f_J^J - f_A^A - f_B^B\right)$

$$\frac{1}{2}\bar{g}^{Al}_{bc}\tilde{t}^{bc}_{Il} = g^{AL}_{CD}\left(2\,T^{BC}_{IL} - T^{BC}_{LI} + \frac{C_0+3C_1}{2}A^{BC}_{IL} + \frac{C_0-3C_1}{2}A^{BC}_{LI}\right), \tag{5.36}$$

$$\begin{aligned} C^{Al}_{a'c}\tilde{t}^{bc}_{Il} = {} & f^{A'}_{B}\, r^{AL}_{A'C}\left(\frac{C_0+3C_1}{2}T^{BC}_{IL} + \frac{C_0-3C_1}{2}T^{BC}_{LI}\right. \\ & \left. + \frac{C_0^2+3C_1^2}{2}A^{BC}_{IL} + \frac{C_0^2-3C_1^2}{2}A^{BC}_{LI}\right) \\ & + f^{A'}_{B}\, r^{AL}_{CB}\left(\frac{C_0+3C_1}{2}T^{CB}_{IL} + \frac{C_0-3C_1}{2}T^{CB}_{LI}\right. \\ & \left. + \frac{C_0^2+3C_1^2}{2}A^{CB}_{IL} + \frac{C_0^2-3C_1^2}{2}A^{CB}_{LI}\right), \end{aligned} \tag{5.37}$$

$$\frac{1}{2}\bar{g}^{kl}_{Mb}\tilde{t}^{Ab}_{kl} = g^{KL}_{MB}\left(2\,T^{AB}_{KL} - T^{BA}_{KL} + \frac{C_0+3C_1}{2}A^{AB}_{KL} + \frac{C_0-3C_1}{2}A^{BA}_{KL}\right), \tag{5.38}$$

$$\begin{aligned} F^{a'}_{M}\bar{R}^{kl}_{a'b}\tilde{t}^{Ab}_{kl} = {} & f^{A'}_{M}\, r^{KL}_{A'B}\left(\frac{C_0+3C_1}{2}T^{AB}_{KL} + \frac{C_0-3C_1}{2}T^{BA}_{KL}\right. \\ & \left. + \frac{C_0^2+3C_1^2}{2}A^{AB}_{KL} + \frac{C_0^2-3C_1^2}{2}A^{BA}_{KL}\right), \end{aligned} \tag{5.39}$$

$$\begin{aligned} F^{a'}_{b}\,\bar{R}^{kl}_{Ma'}\tilde{t}^{Ab}_{kl} = {} & f^{A'}_{B}\, r^{KL}_{MA'}\left(\frac{C_0+3C_1}{2}T^{AB}_{KL} + \frac{C_0-3C_1}{2}T^{BA}_{KL}\right. \\ & \left. + \frac{C_0^2+3C_1^2}{2}A^{AB}_{KL} + \frac{C_0^2-3C_1^2}{2}A^{BA}_{KL}\right), \end{aligned} \tag{5.40}$$

$$\frac{1}{2}\bar{R}^{kl}_{Ac'}\bar{g}^{Mc'}_{kl} = \left(\frac{C_0+3C_1}{2}r^{KL}_{AA'} + \frac{C_0-3C_1}{2}r^{LK}_{AA'}\right)g^{KL}_{MA'}, \tag{5.41}$$

$$\frac{1}{2}\bar{R}^{kl}_{Mc'}\bar{g}^{Ac'}_{kl} = \left(\frac{C_0+3C_1}{2}r^{KL}_{MA'} + \frac{C_0-3C_1}{2}r^{LK}_{MA'}\right)g^{KL}_{AA'}, \tag{5.42}$$

$$\frac{1}{2}\left(F^{k}_{k} + F^{l}_{l}\right)\bar{R}^{kl}_{Mb'}\bar{R}^{Ab'}_{kl} = \left(f^{J}_{K}r^{JL}_{MB} + f^{J}_{L}r^{KJ}_{MB}\right)\left(\frac{C_0^2+3C_1^2}{2}r^{AB'}_{KL} + \frac{C_0^2-3C_1^2}{2}r^{AB'}_{LK}\right), \tag{5.43}$$

$$\frac{1}{2}F^{a'}_{M}\bar{R}^{kl}_{a'b'}\bar{R}^{Ab'}_{kl} = f^{A'}_{M}\, r^{KL}_{A'B}\left(\frac{C_0^2+3C_1^2}{2}r^{AB'}_{KL} + \frac{C_0^2-3C_1^2}{2}r^{AB'}_{LK}\right), \tag{5.44}$$

$$\frac{1}{2}F^{b'}_{a'}\bar{R}^{kl}_{Mc'}\bar{R}^{Ab'}_{kl} = f^{A'}_{B'}r^{KL}_{MA'}\left(\frac{C_0^2+3C_1^2}{2}r^{AB'}_{KL} + \frac{C_0^2-3C_1^2}{2}r^{AB'}_{LK}\right), \tag{5.45}$$

$$\begin{aligned} \frac{1}{2}\bar{R}^{kl}_{Mc'}F^{A}_{\beta}\bar{R}^{\beta c'}_{kl} = {} & r^{KL}_{MC'}\left(\frac{C_0^2+3C_1^2}{2}f^{A}_{B}r^{BC'}_{KL} + \frac{C_0^2-3C_1^2}{2}f^{A}_{B}r^{BC'}_{LK}\right. \\ & \left. + \frac{C_0^2+3C_1^2}{2}f^{A}_{B'}r^{B'C'}_{KL} + \frac{C_0^2-3C_1^2}{2}f^{A}_{B'}r^{B'C'}_{LK}\right). \end{aligned} \tag{5.46}$$

5.5.2 Results for x^2 and y^2 Components of H_2O Quadrupole Moment from F12 Calculations

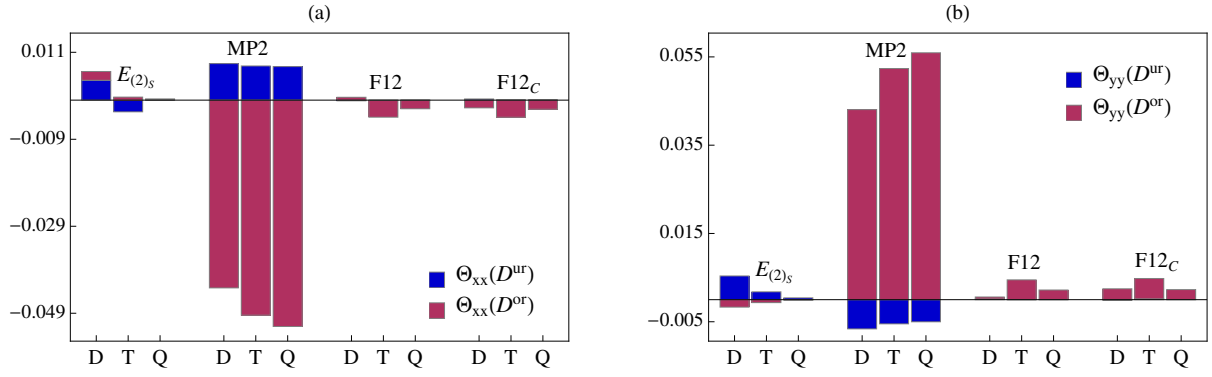

Figure 5.7: The (unrelaxed) one-electron density (D^{ur}) and orbital response (D^{or}) contributions from the CABS singles and correlation corrections to Θ_{xx} and Θ_{yy} for H_2O. F12 and $F12_C$ refer to the F12 correlation corrections without and with the coupling from the MP2 and F12 corrections.

Table 5.8: The derivations of the triple-ζ F12 calculations of Θ_{xx} and Θ_{yy} (a.u.) for H_2O with respect to the estimated basis set limits.[a]

Method	Basis	Θ_{xx}	Θ_{yy}
MP2-F12$_C$ [b]	aug-cc-pVTZ	−0.00065	−0.00223
	cc-pVTZ-F12	0.01130	−0.01023
MP2-F12[b]	aug-cc-pVTZ	−0.00067	−0.00254
	cc-pVTZ-F12	0.01128	−0.01022
		Basis set limit	
MP2	—	−1.84402	1.95386

[a]The basis set limits were the HF/aug-cc-pV6Z values plus the extrapolated MP2 correlation contributions from the aug-cc-pVQZ and aug-cc-pV5Z values. [b]The CABS singles correction was included in the calculations.

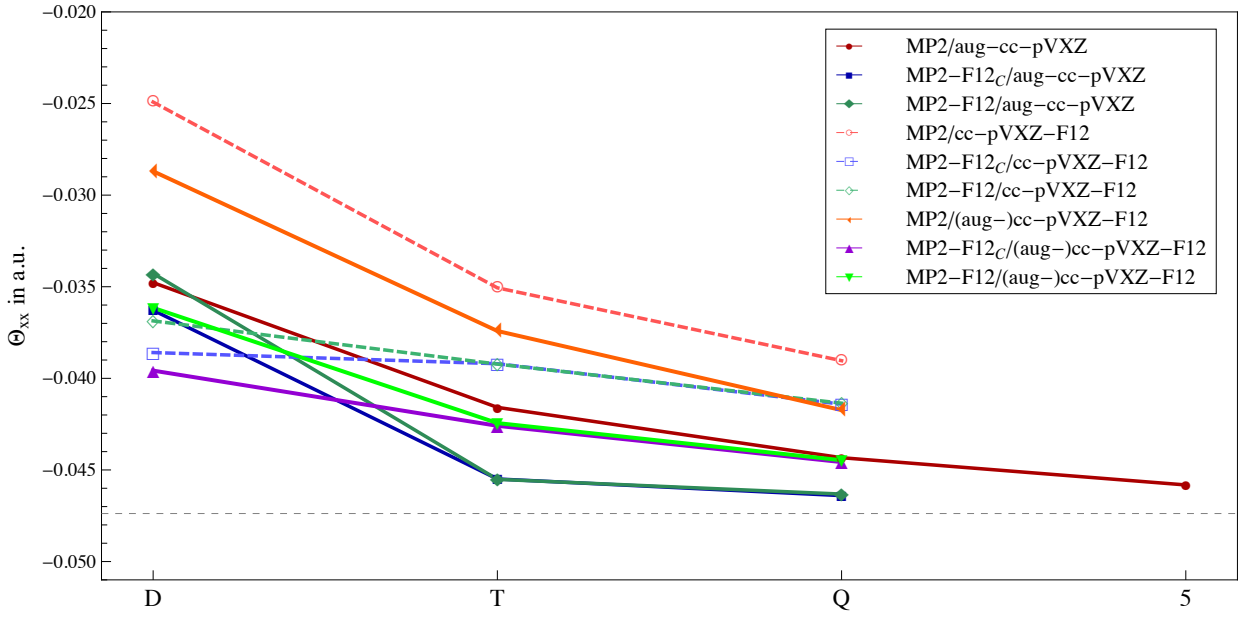

Figure 5.8: Basis set convergence of MP2 correlation contributions to Θ_{xx} for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

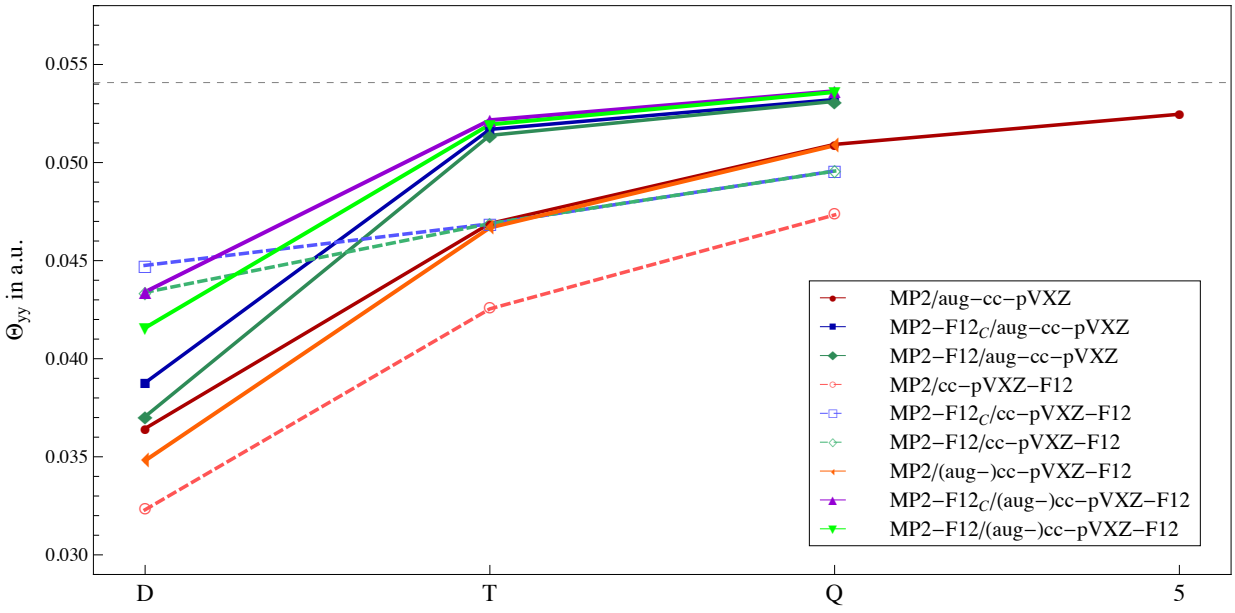

Figure 5.9: Basis set convergences of MP2 correlation contributions to Θ_{yy} for H_2O. MP2-$F12_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

5.5.3 Basis Set Convergence of Various MP2 Correlation Contributions to z Components of Dipole Moment

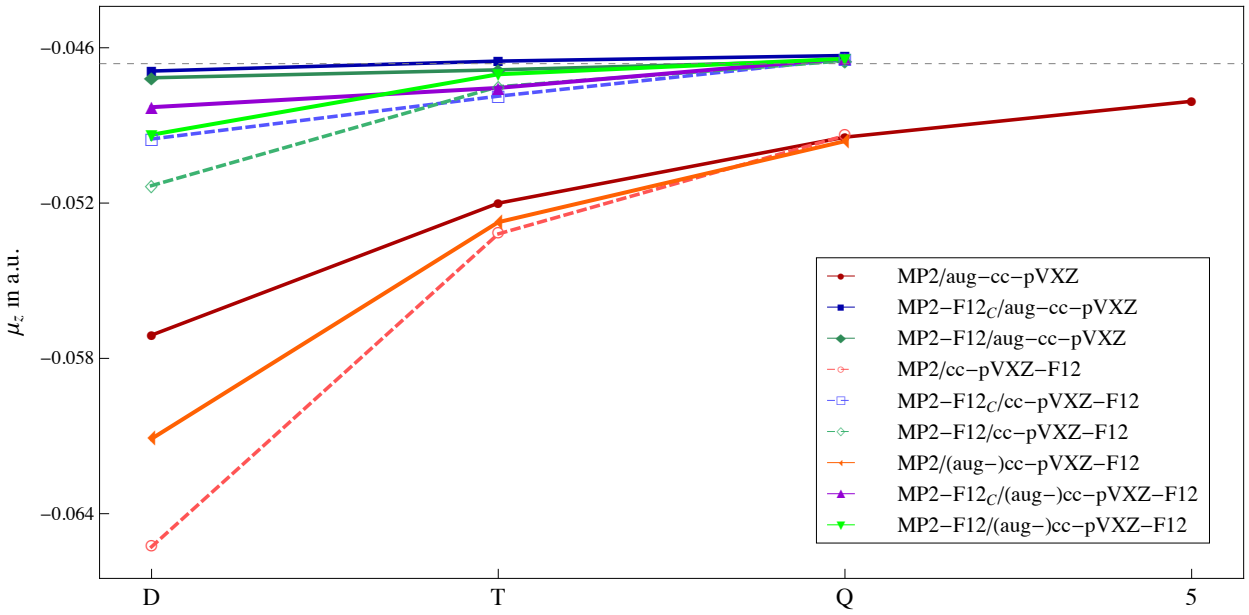

Figure 5.10: Basis set convergence of MP2 correlation contributions to μ_z for BH. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the boron atom while the aug-cc-pVXZ basis sets were used for the hydrogen atom.

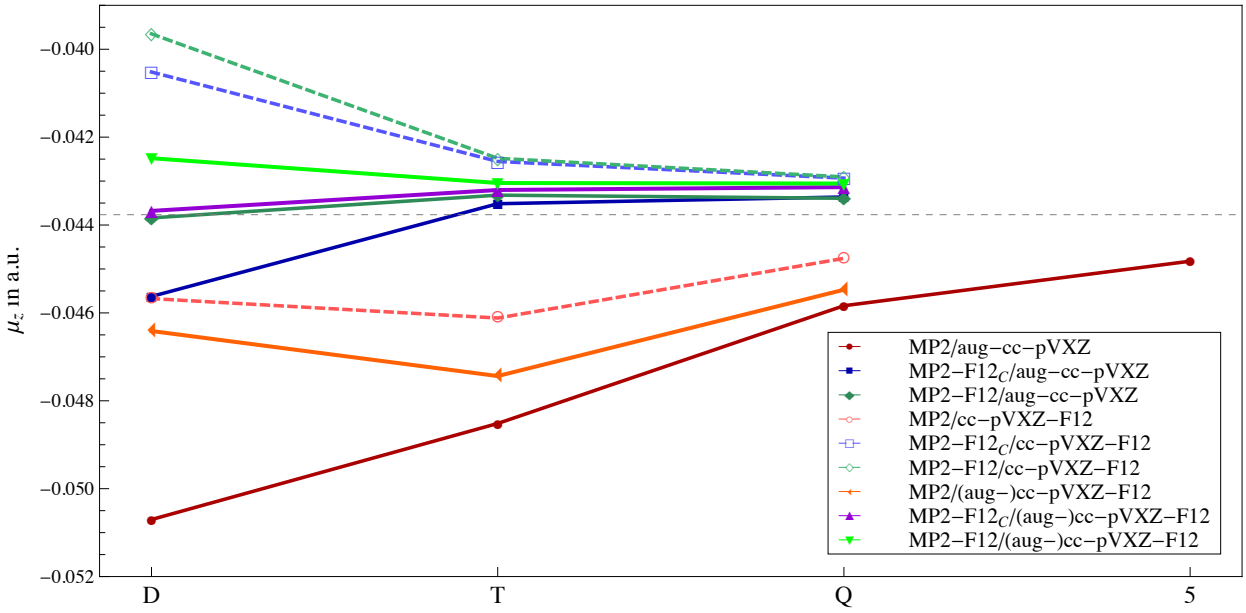

Figure 5.11: Basis set convergence of MP2 correlation contributions to μ_z for the HF molecule. MP2-F12$_{\mathrm{C}}$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the fluorine atom while the aug-cc-pVXZ basis sets were used for the hydrogen atom.

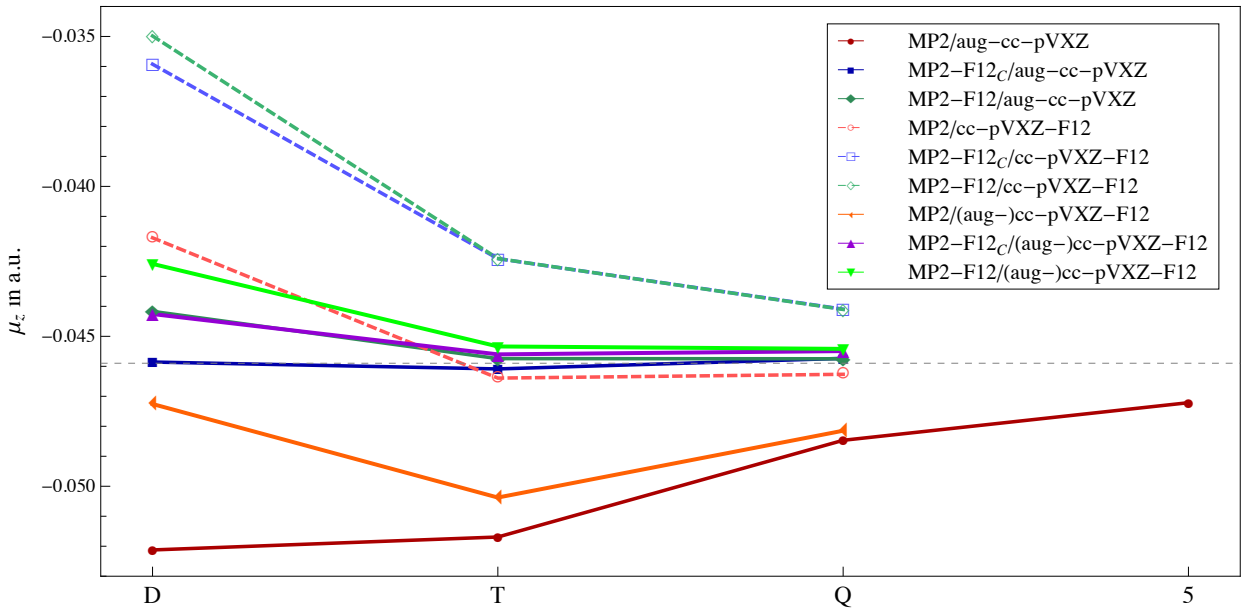

Figure 5.12: Basis set convergence of MP2 correlation contributions to μ_z for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

5.5.4 Basis Set Convergences of Various MP2 Correlation Contributions to z^2 Components of Quadrupole Moment

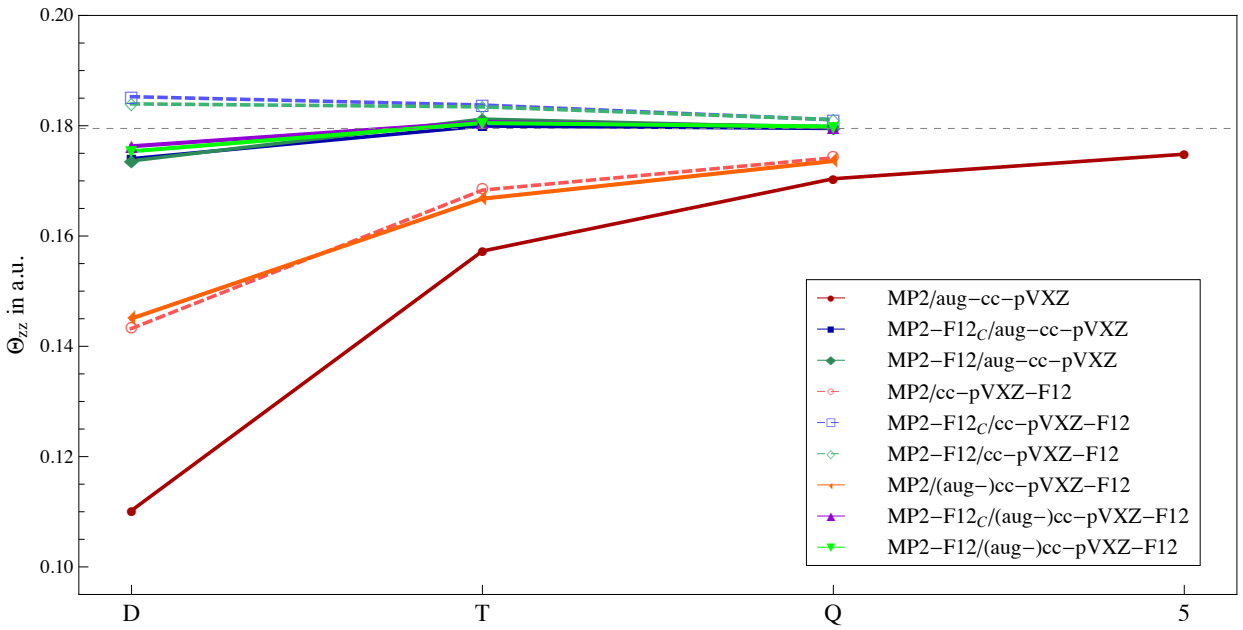

Figure 5.13: Basis set convergences of MP2 correlation contributions to Θ_{zz} for BH. MP2-F12$_{\mathrm{C}}$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the boron atom while the aug-cc-pVXZ basis sets were used for the hydrogen atom.

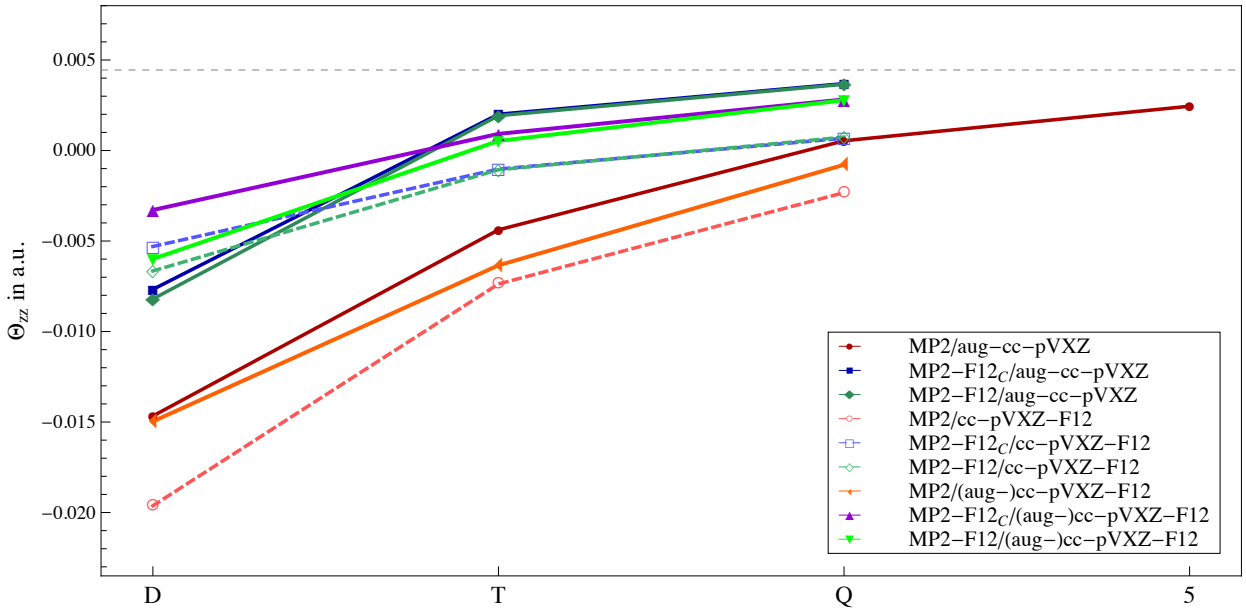

Figure 5.14: Basis set convergences of MP2 correlation contributions to Θ_{zz} for the HF molecule. MP2-F12$_{\mathrm{C}}$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets are used for the fluorine atom while the aug-cc-pVXZ basis sets were used for the hydrogen atom.

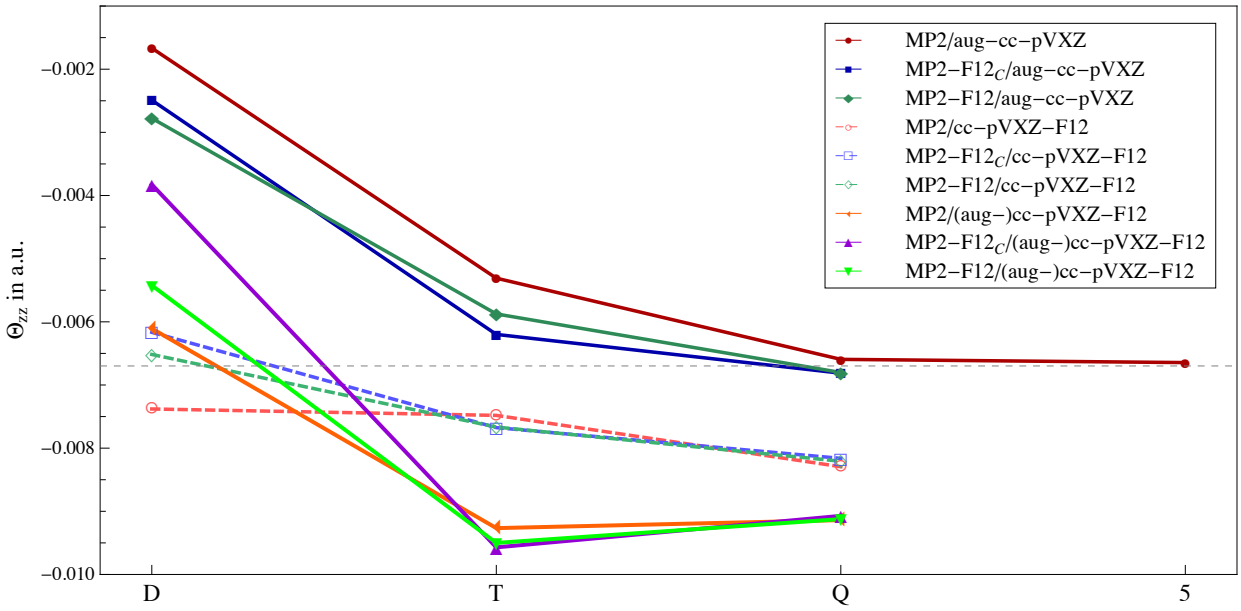

Figure 5.15: Basis set convergence of MP2 correlation contributions to Θ_{zz} for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets are used for the hydrogen atoms.

5.5.5 F12 Corrected Correlation Contributions to the Components of the H_2O Quadrupole Moment, r_α^2, with Different Basis Sets

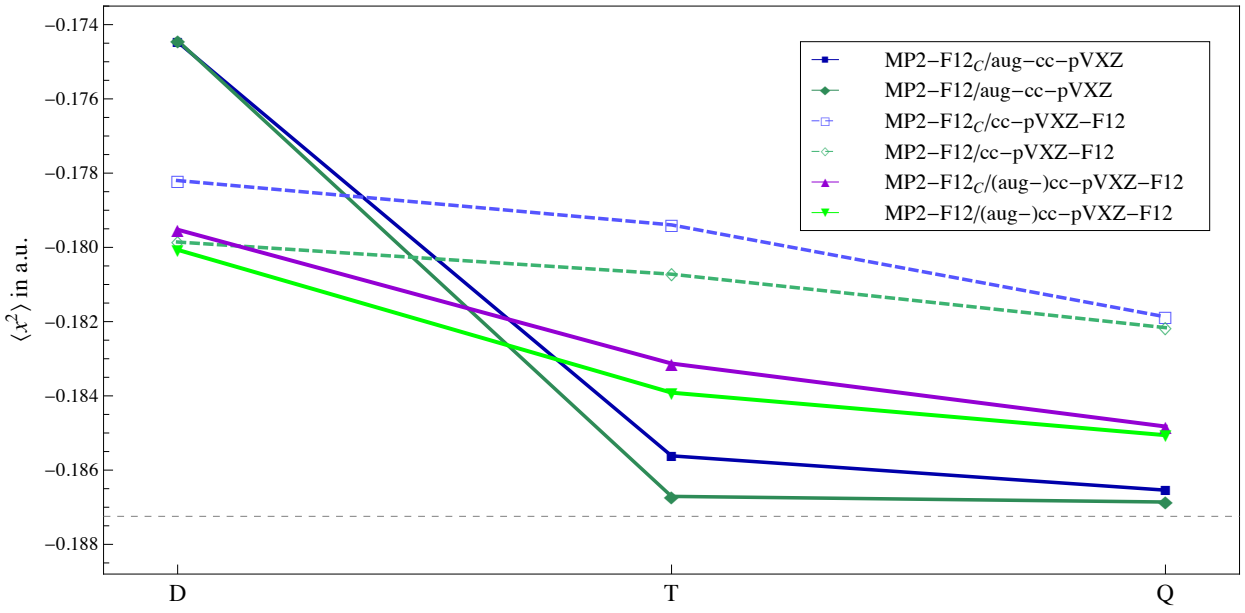

Figure 5.16: Basis set convergence of F12-corrected correlation contributions to $\langle x^2 \rangle$ for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

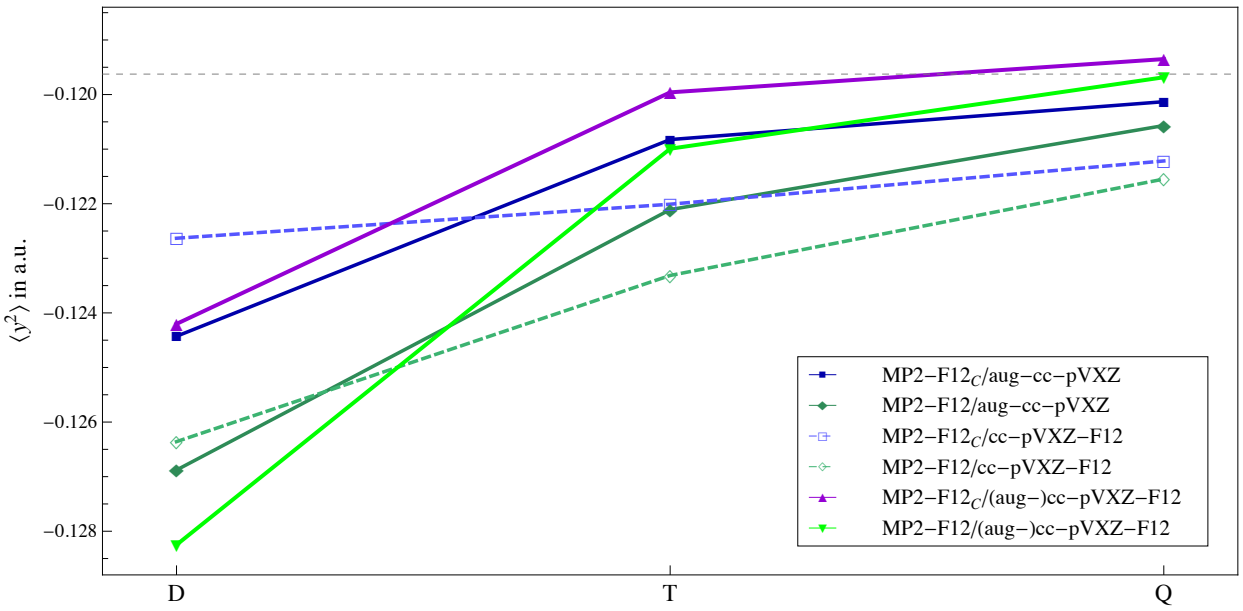

Figure 5.17: Basis set convergence of F12-corrected correlation contributions to $\langle y^2 \rangle$ for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

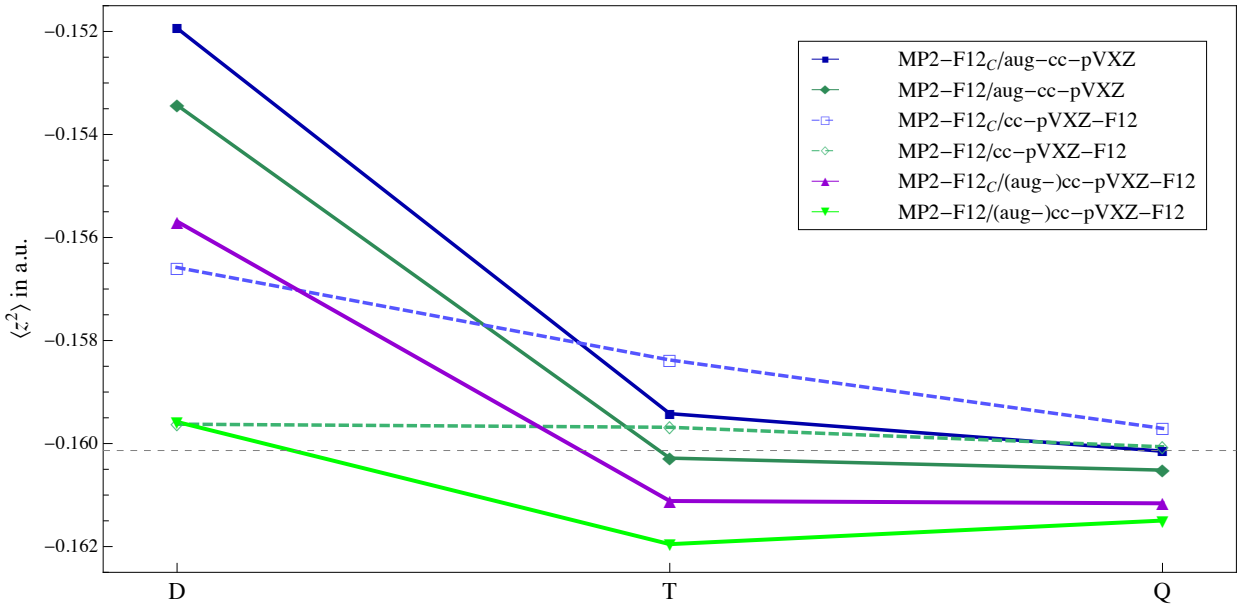

Figure 5.18: Basis set convergence of F12-corrected correlation contributions to $\langle z^2 \rangle$ for H_2O. MP2-F12$_C$ and MP2-F12 refer to the F12-corrected correlation contributions with and without the coupling from the MP2 and F12 corrections. The (aug-)cc-pVXZ-F12 basis sets denote that the cc-pVXZ-F12 basis sets were used for the oxygen atom while the aug-cc-pVXZ basis sets were used for the hydrogen atoms.

Appendix A

Following the convention, we used i, j,..., a, b,..., and p, q,... for the occupied, virtual, and general orbitals in the Hartree-Fock (HF) basis, respectively; α, β,... for the virtual orbitals in the complete basis, and a', b',... for the virtual orbitals in the complete basis that do not belong to the HF basis. In addition, we labeled spin-orbitals with the lowercase indices (i, j, α, β...), the spatial parts of alpha orbitals with uppercase indices (I, J, A, B, ...), and the spatial parts of beta orbitals with uppercase indices with bars ($\bar{I}$, $\bar{J}$, $\bar{A}$, $\bar{B}$, etc.).

Details of tensor notation for second-quantized expressions can be found elsewhere, e.g., Ref. 64. The relationship between tensor and Dirac notations is as follows:

$$O_p^q = \langle p|O(1)|q\rangle \tag{A.1}$$

$$G_{pq}^{rs} = \langle pq|G(1,2)|rs\rangle \tag{A.2}$$

Antisymmettrized matrix elements are denoted with a horizontal line above the operator symbol:

$$\bar{G}_{pq}^{rs} = G_{pq}^{rs} - G_{pq}^{sr}. \tag{A.3}$$

Summation is implied over all pairs of indices that appear in the same term of an expression both in bra and ket (the Einstein summation convention), *unless* the summation is shown explicitly.

Publication List

[1] Zhang, J. and Valeev, E. F., *J. Chem. Theory Comput.*, 2012, **8**, 3175.

[2] Zhang, J. and Valeev, E. F., *Phys. Chem. Chem. Phys.*, 2012, **14**, 7863.

[3] Willow, S. Y.; Zhang, J.; Valeev, E. F. and Hirata, S., *J. Chem. Phys.*, 2014, **140**, 031101.

Bibliography

[1] Helgaker, T.; Jørgensen, P. and Olsen, J., *Modern Electronic Structure Theory*, Wiley, Chichester, first ed., 2000.

[2] Kong, L.; Bischoff, F. A. and Valeev, E. F., *Chem. Rev.*, 2012, **112**, 75.

[3] Tew, D. P. and Klopper, W., *J. Chem. Phys.*, 2005, **123**, 074101.

[4] Köhn, A. and Tew, D. P., *J. Chem. Phys.*, 2010, **133**, 174117.

[5] Zheng, J.; Zhao, Y. and Truhlar, D. G., *J. Chem. Theory Comput.*, 2007, **3**, 569.

[6] Karton, A.; Tarnopolsky, A.; Lamere, J.-F.; Schatz, G. C. and Martin, J. M. L., *J. Phys. Chem. A*, 2008, **112**, 12868.

[7] Zheng, J.; Zhao, Y. and Truhlar, D. G., *J. Chem. Theory Comput.*, 2009, **5**, 808.

[8] Kutzelnigg, W. and Morgan III, J. D., *J. Chem. Phys.*, 1992, **96**, 4484.

[9] Kutzelnigg, W., *Theor. Chim. Acta*, 1985, **68**, 445.

[10] Tew, D. P.; Hättig, C.; Bachorz, R. A. and Klopper, W. In *Recent Progress in Coupled Cluster Methods*, Čársky, P.; Paldus, J. and Pittner, J., Eds.; Springer, New York, 2010; page 535.

[11] Tunega, D.; Noga, J. and Klopper, W., *Chem. Phys. Lett.*, 1997, **269**, 435.

[12] Kordel, E.; Villani, C. and Klopper, W., *J. Chem. Phys.*, 2005, **122**, 214306.

[13] Höfener, S.; Hättig, C. and Klopper, W., *Z. Phys. Chem.*, 2010, **224**, 695.

[14] Yang, J. and Hättig, C., *J. Chem. Phys.*, 2009, **131**, 074102.

[15] Peterson, K. A.; Adler, T. B. and Werner, H.-J., *J. Chem. Phys.*, 2008, **128**, 084102.

[16] Yousaf, K. E. and Peterson, K. A., *J. Chem. Phys.*, 2008, **129**, 184108.

[17] Peterson, K. A.; Krause, C.; Stoll, H.; Hill, J. G. and Werner, H.-J., *Mol. Phys.*, 2011, **109**, 2607.

[18] Dunning, T. H., *J. Chem. Phys.*, 1989, **90**, 1007.

[19] Kendall, R. A.; Dunning, T. H. and Harrison, R. J., *J. Chem. Phys.*, 1992, **96**, 6796.

[20] Woon, D. E. and Dunning, T. H., *J. Chem. Phys.*, 1993, **98**, 1358.

[21] Born, M. and Oppenheimer, J. R., *Ann. Physik*, 1927, **84**, 457.

[22] Szabo, A. and Ostlund, N. S., *Modern Quantum Chemistry : Introduction to Advanced Electronic Structure Theory*, Dover Publications, Inc., Mineola, N.Y., 1996.

[23] Roothaan, C. C. J., *Rev. Mod. Phys.*, 1951, **23**, 69.

[24] Raghavachari, K. and Anderson, J. B., *J. Phys. Chem.*, 2012, **100**, 12960.

[25] Bartlett, R. J. and Stanton, J. F. In *Reviews in Computational Chemistry Volume V*, Lipkowitz, K. B. and Boyd, D. B., Eds., Reviews in Computational Chemistry; VCH, New York, 1994; chapter 2, page 65.

[26] Cremer, D., *WIREs Comput. Mol. Sci.*, 2013, **3**, 482.

[27] Crawford, T. D. and Schaefer III, H. F. In *Reviews in Computational Chemistry Volume 14*, Lipkowitz, K. B. and Boyd, D. B., Eds., number 14 in Reviews in Computational Chemistry; WILEY-VCH, New York, 2000; chapter 2, page 33.

[28] Bartlett, R. and Musiał, M., *Rev. Mod. Phys.*, 2007, **79**, 291.

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

[30] Hohenberg, P. and Kohn, W., *Phys. Rev.*, 1964, **136**, B864.

[31] Kohn, W. and Sham, L., *J. Phys. Rev. A*, 1965, **140**, 1133.

[32] Cremer, D., *Mol. Phys.*, 2001, **99**, 1899.

[33] Helgaker, T.; Klopper, W.; Koch, H. and Noga, J., *J. Chem. Phys.*, 1997, **106**, 9639.

[34] Pack, R. T. and Brown, W., *J. Chem. Phys.*, 1966, **45**, 556.

[35] Hylleraas, E. A., *Z. Phys.*, 1929, **54**, 347.

[36] Coulson, C. A. and Neilson, A. H., *Proc. Phys. Soc.*, 1961, **78**, 831.

[37] Boys, S. F., *Proc. Roy. Soc. (London) A*, 1960, **258**, 402.

[38] Singer, K., *Proc. Roy. Soc. (London) A*, 1960, **258**, 412.

[39] Boys, S. F. and Handy, N. C., *Proc. Roy. Soc. (London) A*, 1969, **309**, 209.

[40] Boys, S. F. and Handy, N. C., *Proc. Roy. Soc. (London) A*, 1969, **310**, 43.

[41] Röhse, R.; Klopper, W. and Kutzelnigg, W., *J. Chem. Phys.*, 1993, **99**, 8830.

[42] Clary, D. C. and Handy, N. C., *Phys. Rev. A*, 1976, **14**, 1607.

[43] James, H. M. and Coolidge, A. S., *J. Chem. Phys.*, 1933, **1**, 825.

[44] Kołos, W. and Wolniewicz, L., *J. Chem. Phys.*, 1965, **43**, 2429.

[45] Jeziorski, B. and Szalewicz, K., *Phys. Rev. A*, 1979, **19**, 2360.

[46] Szalewicz, K. and Jeziorski, B., *Mol. Phys.*, 1979, **38**, 191.

[47] Szalewicz, K.; Jeziorski, B. and Monkhorst, H. J., *Chem. Phys. Lett.*, 1982, **91**, 169.

[48] Szalewicz, K.; Zabolitzky, J. G.; Jeziorski, B. and Monkhorst, H. J., *J. Chem. Phys.*, 1984, **81**, 2723.

[49] Wenzel, K. B.; Zabolitzky, J. G.; Szalewicz, K.; Jeziorski, B. and Monkhorst, H. J., *J. Chem. Phys.*, 1986, **85**, 3964.

[50] Bukowski, R.; Jeziorski, B.; Rybak, S. and Szalewicz, K., *J. Chem. Phys.*, 1995, **102**, 888.

[51] Bukowski, R.; Jeziorski, B. and Szalewicz, K., *J. Chem. Phys.*, 1999, **110**, 4165.

[52] Kozlowski, P. M. and Adamowicz, L., *J. Chem. Phys.*, 1991, **95**, 6681.

[53] Cencek, W. and Rychlewski, J., *J. Chem. Phys.*, 1993, **98**, 1252.

[54] Cencek, W. and Rychlewski, J., *J. Chem. Phys.*, 1995, **102**, 2533.

[55] Cencek, W. and Rychlewski, J., *Chem. Phys. Lett.*, 2000, **320**, 549.

[56] Ten-no, S., *Chem. Phys. Lett.*, 2000, **330**, 169.

[57] Luo, H., *J. Chem. Phys.*, 2010, **133**, 154109.

[58] Klopper, W., *Chem. Phys. Lett.*, 1991, **186**, 583.

[59] Noga, J.; Kutzelnigg, W. and Klopper, W., *Chem. Phys. Lett.*, 1992, **199**, 497.

[60] Noga, J. and Kutzelnigg, W., *J. Chem. Phys.*, 1994, **101**, 7738.

[61] Gdanitz, R. J., *Chem. Phys. Lett.*, 1993, **210**, 253.

[62] Gdanitz, R. J., *Chem. Phys. Lett.*, 1998, **283**, 253.

[63] Klopper, W.; Manby, F. R.; Ten-no, S. and Valeev, E. F., *Int. Rev. Phys. Chem.*, 2006, **25**, 427.

[64] Kutzelnigg, W. and Klopper, W., *J. Chem. Phys.*, 1991, **94**, 1985.

[65] Klopper, W. and Samson, C. C. M., *J. Chem. Phys.*, 2002, **116**, 6397.

[66] Valeev, E. F., *Chem. Phys. Lett.*, 2004, **395**, 190.

[67] Kedžuch, S.; Milko, M. and Noga, J., *Int. J. Quant. Chem.*, 2005, **105**, 929.

[68] Ten-no, S., *J. Chem. Phys.*, 2004, **121**, 117.

[69] Ten-no, S., *J. Chem. Phys.*, 2007, **126**, 014108.

[70] Bokhan, D.; Bernadotte, S. and Ten-no, S., *Chem. Phys. Lett.*, 2009, **469**, 214.

[71] Ten-no, S., *Chem. Phys. Lett.*, 2004, **398**, 56.

[72] Persson, B. J. and Taylor, P. R., *J. Chem. Phys.*, 1996, **105**, 5915.

[73] May, A. J.; Valeev, E. F.; Polly, R. and Manby, F. R., *Phys. Chem. Chem. Phys.*, 2005, **7**, 2710.

[74] Bachorz, R. A.; Klopper, W. and Gutowski, M., *J. Chem. Phys.*, 2007, **126**, 085101.

[75] Ten-no, S., *Chem. Phys. Lett.*, 2007, **447**, 175.

[76] Knizia, G. and Werner, H.-J., *J. Chem. Phys.*, 2008, **128**, 154103.

[77] Kordel, E.; Villani, C. and Klopper, W., *Mol. Phys.*, 2007, **105**, 2565.

[78] Manby, F. R.; Werner, H.-J.; Adler, T. B. and May, A. J., *J. Chem. Phys.*, 2006, **124**, 094103.

[79] Valeev, E. F., *J. Chem. Phys.*, 2006, **125**, 244106.

[80] Adler, T. B.; Knizia, G. and Werner, H.-J., *J. Chem. Phys.*, 2007, **127**, 221106.

[81] Noga, J.; Kedžuch, S. and Šimunek, J., *J. Chem. Phys.*, 2007, **127**, 034106.

[82] Hill, J. G. and Peterson, K. A., *Phys. Chem. Chem. Phys.*, 2010, **12**, 10460.

[83] Hill, J. G.; Mazumder, S. and Peterson, K. A., *J. Chem. Phys.*, 2010, **132**, 054108.

[84] Yousaf, K. E. and Peterson, K. A., *Chem. Phys. Lett.*, 2009, **476**, 303.

[85] Bischoff, F. A.; Wolfsegger, S.; Tew, D. P. and Klopper, W., *Mol. Phys.*, 2009, **107**, 963.

[86] Noga, J.; Valiron, P. and Klopper, W., *J. Chem. Phys.*, 2001, **115**, 2022.

[87] Klopper, W. and Noga, J., *ChemPhysChem*, 2003, **4**, 32.

[88] Shiozaki, T.; Kamiya, M.; Hirata, S. and Valeev, E. F., *J. Chem. Phys.*, 2008, **129**, 071101.

[89] Shiozaki, T.; Kamiya, M.; Hirata, S. and Valeev, E. F., *Phys. Chem. Chem. Phys.*, 2008, **10**, 3358.

[90] Shiozaki, T.; Kamiya, M.; Hirata, S. and Valeev, E. F., *J. Chem. Phys.*, 2009, **130**, 054101.

[91] Köhn, A.; Richings, G. W. and Tew, D. P., *J. Chem. Phys.*, 2008, **129**, 201103.

[92] Tew, D.; Klopper, W. and ttig, C. H., *Chem. Phys. Lett.*, 2008, **452**, 326.

[93] Zhang, J. and Valeev, E. F., *J. Chem. Theory Comput.*, 2012, **8**, 3175.

[94] Fliegl, H.; Klopper, W. and Hättig, C., *J. Chem. Phys.*, 2005, **122**, 084107.

[95] Hättig, C.; Tew, D. P. and Köhn, A., *J. Chem. Phys.*, 2010, **132**, 231102.

[96] Knizia, G.; Adler, T. B. and Werner, H.-J., *J. Chem. Phys.*, 2009, **130**, 054104.

[97] Valeev, E. F., *Phys. Chem. Chem. Phys.*, 2008, **10**, 106.

[98] Torheyden, M. and Valeev, E. F., *Phys. Chem. Chem. Phys.*, 2008, **10**, 3410.

[99] Angeli, C.; Bak, K. L.; Bakken, V.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Hättig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Helgaker, T.; Hettema, H.; Jensen, H. J. A.; Jonsson, D.; Jørgensen, P.; Kirpekar, S.; Klopper, W.; Kobayashi, R.; Koch, H.; Ligabue, A.; s., O. B. L.; Mikkelsen, K. V.; Norman, P.; Olsen, J.; Packer, M. J.; Pedersen, T. B.; Rinkevicius, Z.; Rudberg, E.; Ruden, T. A.; Ruud, K.; Salek, P.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O.; Wilson, D. J.; and Ågren, H., Dalton, a molecular electronic structure program, release 2.0, 2005, see: http:// www.kjemi.uio.no/software/dalton/dalton.html.

[100] TURBOMOLE V6.5 2013, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

[101] Janssen, C. L.; Nielsen, I. B.; Leininger, M. L.; Valeev, E. F.; Kenny, J. P. and Seidl, E. T., The massively parallel quantum chemistry program (mpqc): Version 3.0 alpha.

[102] Tew, D. P.; Klopper, W.; Neiss, C. and H ttig, C., *Phys. Chem. Chem. Phys.*, 2007, **9**, 1921.

[103] Tew, D. P.; Klopper, W.; Neiss, C. and Hättig, C., *Phys. Chem. Chem. Phys.*, 2008, **10**, 6325.

[104] Fliegl, H.; Hättig, C. and Klopper, W., *Int. J. Quantum Chem.*, 2006, **106**, 2306.

[105] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. and Wang, M., Molpro, version 2012.1, a package of ab initio programs, 2012.

[106] Werner, H.-J.; Adler, T. B.; Knizia, G. and Manby, F. R. In *Recent Progress in Coupled Cluster Methods*, Čársky, P.; Paldus, J. and Pittner, J., Eds.; Springer, 2010.

[107] Stanton, J. F., *Chem. Phys. Lett.*, 1997, **281**, 130.

[108] Löwdin, P.-O., *J. Math. Phys.*, 1962, **3**, 969.

[109] Valeev, E. F. and Crawford, T. D., *J. Chem. Phys.*, 2008, **128**, 244113.

[110] Torheyden, M. and Valeev, E. F., *J. Chem. Phys.*, 2009, **131**, 171103.

[111] Shiozaki, T. and Werner, H.-J., *J. Chem. Phys.*, 2010, **133**, 141103.

[112] Shiozaki, T.; Knizia, G. and Werner, H.-J., *J. Chem. Phys.*, 2011, **134**, 034113.

[113] Willow, S. Y.; Zhang, J.; Valeev, E. F. and Hirata, S., *J. Chem. Phys.*, 2014, **140**, 031101.

[114] Willow, S. Y.; Kim, K. S. and Hirata, S., *J. Chem. Phys.*, 2012, **137**, 204122.

[115] Luo, H.; Hackbusch, W. and Flad, H.-J., *Mol. Phys.*, 2010, **108**, 425.

[116] Chinnamsetty, S. R.; Luo, H.; Hackbusch, W.; Flad, H.-J. and Uschmajew, A., *Chem. Phys.*, 2012, **401**, 36.

[117] Klopper, W. and Kutzelnigg, W., *Chemical Physics Letters*, 1987, **134**, 17.

[118] Manby, F. R., *J. Chem. Phys.*, 2003, **119**, 4607.

[119] Ten-no, S. and Manby, F. R., *J. Chem. Phys.*, 2003, **119**, 5358.

[120] Huang, X.; Valeev, E. F. and Lee, T. J., *J. Chem. Phys.*, 2010, **133**, 244108.

[121] Neiss, C. and Hättig, C., *J. Chem. Phys.*, 2007, **126**, 154101.

[122] Bischoff, F. and Klopper, W., *J. Chem. Phys.*, 2010, **132**, 094108.

[123] Bischoff, F.; Valeev, E.; Klopper, W. and Janssen, C., *J. Chem. Phys.*, 2010, **132**, 214104.

[124] Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J. and Stanton, J. F., *J. Chem. Phys.*, 2008, **128**, 114111.

[125] Neese, F. and Valeev, E. F., *J. Chem. Theory Comput.*, 2011, **7**, 33.

[126] Tajti, A.; Szalay, P.; Császár, A.; Kállay, M.; Gauss, J.; Valeev, E.; Flowers, B.; Vázquez, J. and Stanton, J., *J. Chem. Phys.*, 2004, **121**, 11599.

[127] Bokhan, D.; Bernadotte, S. and Ten-no, S., *J. Chem. Phys.*, 2009, **131**, 084105.

[128] Dunning, T. H.; Peterson, K. A. and Wilson, A. K., *J. Chem. Phys.*, 2001, **114**, 9244.

[129] Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P. and Allen, W. D., *J. Comput. Chem.*, 2007, **28**, 1610.

[130] Schwenke, D. W., *J. Chem. Phys.*, 2005, **122**, 014107.

[131] Hobson, S. L.; Valeev, E. F.; Császár, A. G. and Stanton, J. F., *Mol. Phys.*, 2009, **107**, 1153.

[132] Halkier, A.; Larsen, H.; Olsen, J.; Jorgensen, P. and Gauss, J., *J. Chem. Phys.*, 1999, **110**, 734.

[133] Peterson, K. A. and Dunning, T. H., *J. Mol. Struct. (THEOCHEM)*, 1997, **400**, 93.

[134] Gauss, J. and Cremer, D., *Adv. Quantum Chem.*, 1992, **23**, 205.

[135] Buckingham, A. D., *Adv. Chem. Phys.*, 1967, **12**, 107.