# 國 立 交 通 大 學

## 應用化學系分子科學碩士班
## 碩士論文

H$_2$–HCl 凡德瓦錯合物之全初始化互相作用能曲面

An *ab initio* interaction energy surface of the H$_2$–HCl van der Waals complex

研究生：李彥廷

指導教授：魏恆理 教授

中華民國一百零三年四月

H$_2$–HCl 凡德瓦錯合物之全初始化互相作用能曲面

An *ab initio* interaction energy surface of the H$_2$–HCl van der Waals complex

研 究 生：李彥廷　　　　　　　　　　　　　Student：Yen-Ting Li

指導教授：魏恆理　　　　　　　　　　　　　Advisor：Henryk Witek

國 立 交 通 大 學

應 用 化 學 系 分 子 科 學 碩 士 班

碩 士 論 文

A Thesis

Submitted to M. S. Program in Molecular Science,

Department of Applied Chemistry

National Chiao Tung University

In partial Fulfillment of the Requirements

For the Degree of

Master

In

Molecular Science

April 2014

Hsinchu, Taiwan, Republic of China

中 華 民 國 一 百 零 三 年 四 月

# H$_2$–HCl 凡德瓦錯合物之全初始化互相作用能曲面

研究生:李彥廷　　　　　　　　　　　　指導教授:魏恆理

國立交通大學應用化學系分子科學碩士班

# 摘要

一個六維 H$_2$-HCl 錯合物之分子間互相作用能曲面已被提出。全初始化計算點完成於六維的網格,包含了 H-Cl 及 H-H 距離的相關。首先,利用高效率明確相關耦合簇超分子方法(efficient explicitly correlated coupled-cluster supermolecular method (CCSD(T)-F12b)) 及大基底函數組,通過外插法計算出互相作用能。接著,三重迭代及四重非迭代激發貢獻,由較小的基底函數組計算得出並且加入。此研究中也包含了內層電子互相作用及相對論修正。所得到的互相作用能接著對 HCl 及 H$_2$ 之基態振動波函數進行平均。

由此進行解析擬合,產生四維勢能面。在勢能井,對應到真實完整基底函數組的不確定性小於 0.6 波數。藉由我們提出的兩步擬合,勢能面的形貌及能量皆能正確地呈現。擬合步驟在勢能井最大誤差及平均誤差分別為 $0.47\mathrm{cm}^{-1}$ 及 $0.0066\ \mathrm{cm}^{-1}$。擬合使用簡單的三角函數。這些簡單的三角函數在未來工作中容易進行操作。相似的互相作用能曲面可以依照本論文中所發展的方法來產生。

# An *ab initio* interaction energy surface of the H$_2$–HCl van der Waals complex

Student: Yen-Ting Li                    Advisor: Henryk Witek

M. S. Program in Molecular Science,

Department of Applied Chemistry

National Chiao Tung University

## Abstract

A six-dimensional intermolecular potential energy surfaces for the H$_2$–HCl complex is presented. The *ab initio* points have been computed on a six-dimensional grid including the dependence on the H–Cl and H–H separations. The interaction energies were first calculated using the efficient explicitly correlated coupled-cluster supermolecular method (CCSD(T)-F12b) with large basis sets, followed by an extrapolation procedure. Next, a contribution from iterative triple and noniterative quadruple excitations was added from calculations in smaller basis sets. The core electron correlation and relativistic correction were also included. The resulting interaction energies were then averaged over the vibrational ground-state wave functions of HCl and H$_2$.

The final four-dimensional potential energy surfaces were fitted by analytic expressions. In the potential well, the uncertainty corresponding to the true complete basis set limit is less than 0.6 cm$^{-1}$. We proposed a two-step fitting procedure which reproduced well both the correct shape of the whole potential energy surface and interaction energy itself. The maximum error of the fitting in the attractive part of potential is 0.47 cm$^{-1}$ and the mean error is 0.0066 cm$^{-1}$. The simple trigonometric

functions were used which are easy to manipulate in future work. Similar interaction

energy surface can be produced based on methodology developed in presented Thesis.

關鍵詞/Keywords:互相作用能(Interaction energy)、勢能面(Potential energy surface)、

凡德瓦錯合物(van der Waals' complexes)
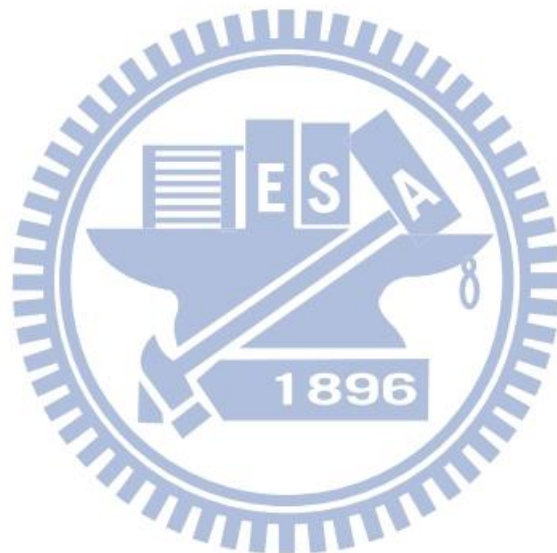
# Acknowledgment

First, I would like to sincerely express my appreciation to my advisor, Henryk, who gave me an opportunity to join this outstanding group. He gave me a lot of precious guides in research also in life. His kind, witty personality and enthusiasm about science impressed on me.

I would like to thank Prof. Zhu and Prof. Shigeto, member the defense committee, for taking time to read my thesis and their suggestions.

I would like to thank Michał for his kind guidance and assistance. Discussions with him have been insightful, and I would never be able to finish this thesis without his help. I am treasure the time with his family, Lucas and Ola.

I would like to thank my friends in this group: Chien-Pin, Bing-hou, Wen-yang Chin-chai, and also the former group members Shu-Wei and Po-Yu. They helped me in many aspects and enrich my daily life.
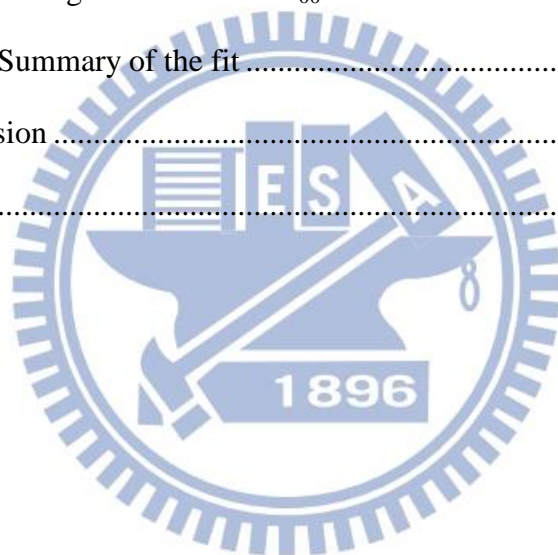
Thank the Lord for giving me these treasurable times. The brothers and sisters in church always take care of me like family and make me feel warm. Their prayers always give me strength to face the difficulty.

Finally, I would like to express the deepest gratitude to my dear parents and my wife, who support me both mentally and economically. Because of them, I could pay all attention on study and through this stage in my life.

# Contents

# List of Figures

VIII

# List of Tables

XI

# Chapter 1

# Introduction

The importance of intermolecular forces in Nature is very difficult to overestimate. It is sufficient to say that the existence of liquids and solid is due to intermolecular interactions. In the absence of intermolecular interactions our world would be a uniform ideal gas.

A detailed knowledge of intermolecular interactions is required to solve a wide class of problems in physics, chemistry and biology. The thermodynamic properties of gases and liquids and their kinetic characteristics are determined by the nature of intermolecular interactions. Intermolecular forces also determine to a large degree the properties of crystals, such as the equilibrium geometry, the binding energy, etc.

Intermolecular interactions are involved in the formation of chemical complexes, such as charge-transfer and hydrogen-bond complexes. Study of the mechanism of elementary chemical reactions is impossible without the knowledge of the exchange processes between the translational and electron-vibration energies, which depend on the interaction of particles under collisions. Knowledge of the potential surface, characterizing the mutual trajectories of the reactants, is necessary to obtain the rates of chemical reactions.

Theoretical modeling of these properties requires knowledge of the usually multidimensional potential energy surface for a wide range of separations, which have to be found independently. Owing to the fast development of computing technology and enhancement of the speed of the central processing units (CPUs), we are able to

calculate the full-dimensional potential energy surface for $H_2$–HCl with errors below one wavenumber.

In 2004, Jankowski proposed[1] a method for the generation of highly accurate, full-dimensional interaction energy surfaces for weakly interacting subsystems (Ar–HF). The method is based on the local expansion of the exact interaction energy surface in Taylor series with respect to the intramolecular coordinates. This leads to significant savings in computational time for the full-dimensional interaction energy surfaces. Later in 2005, Jankowski and Szalewicz[2] applied this method on $H_2$–CO. The ab initio points have been computed on a five-dimensional grid including the dependence on H–H separation (the C–O separation was fixed). The surface has then been obtained by averaging over the intramolecular vibration of $H_2$. In 2013, Szalewicz *et al.* presented[3] a new *ab initio* interaction potential energy surfaces for the $H_2$–CO complex computed on a six-dimensional grid. The resulting interaction energies were then averaged over the ground–state and both ground– and first–excited–states vibrational wave functions of $H_2$ and CO, respectively. Theoretical infrared spectra calculated from these surfaces agree extremely well, to within a few hundredth of a wavenumber, with the experimental spectra of *para* and *ortho* $H_2$–CO complexes. In the latter case, this agreement enabled an assignment of the experimental spectrum, ten years after it had been measured.

In 2002, Anderson *et al.*[4] studied the behavior of HCl and DCl isolated in solid *para*hydrogen (p$H_2$). They found that the gas phase vibrational and rotational quantum numbers of the dopant are conserved within the p$H_2$ solid. In addition, the pure vibrational $Q_1(0)$ ($v$=1←0, $J$=0←0) $H_2$ transition, which is infrared inactive in pure solid p$H_2$, is detected in the HCl doped sample. They propose that this transition is induced in p$H_2$ molecules by neighboring HCl molecules through a weak "overlap

induction" mechanism that is the only induction mechanism operative for $J$=0 impurities in pH$_2$. Rovibrational transitions are also detected near the induced $Q_1(0)$ H$_2$ absorption.

In this Thesis, we tried to provide an accurate, vibrationally averaged H$_2$–HCl interaction energy surface that can be used for simulating the spectrum of HCl trapped in pH$_2$. Furthermore, analogous methodology can be used for generating also other H$_2$–HX (X=F, Br) interaction energy surfaces. The detailed introduction to the computational method used in this Thesis is given in chapter 2. The definition of the coordinate system, grid points, and fitting functions are given in chapter 3. The investigation of the basis set convergence for different methods and features of potential surface are discussed in chapter 4. The conclusions of this Thesis is given in chapter 5.

# 1.1 The concept of interatomic potential and adiabatic approximation

A consistent theory of intermolecular forces can only be developed on the basis of quantum-mechanical principles. Because of the quantum nature of the electronic and nuclear motions, the solution of the intermolecular interaction problem reduces to solving the Schrödinger equation for a system of interacting molecules. Such a problem may be solved only after adopting some approximations. A substantial simplification may be achieved because of the possibility of separating the electronic and nuclear subsystems and introducing the concept of the adiabatic potential. This approach, denoted as the adiabatic approximation, is based on a large difference between the masses of electrons and nuclei.

Within the adiabatic approach, the electron subsystem is studied with fixed nuclei positions. In the Schrödinger equation, the nuclear kinetic energy operator is neglected and the nuclear coordinates are treated as fixed parameters. The Schrödinger equation splits into two: one for the electron motion with fixed nuclei and the other for the nuclear motion with the electron energy as the potential energy.

Below, I represent the basic equations for the system of two atoms A and B with $N_A$ and $N_B$ electrons, respectively, and the total number of electrons in the system N= $N_{A+}$ $N_B$. Let us denote the set of 3N electron coordinates as $r$ and the distance between nuclei as $R$. The total wave function in the adiabatic approximation is written as a simple product:

$$\Psi_{mv}(r,R) = \chi_{mv}(R)\Psi_m(r,R) \tag{1.1}$$

where we denote the wave function describing the nuclear motions as $\chi_{mv}(R)$ and the wave function of the N-electron system in a quantum state $m$ as $\Psi_m(r,R)$: for each electronic quantum state $m$ there is a corresponding set of nuclear quantum states $v$ .

The electronic wave function $\Psi_m(r,R)$ has to satisfy the Schrödinger equation for the electronic motion:

$$H_e\Psi_m(r,R) = E_m(R)\Psi_m(r,R) \tag{1.2}$$

with the Hamiltonian:

$$H_e = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}(\frac{Z_a e^2}{r_{ai}} + \frac{Z_b e^2}{r_{bi}}) + \sum_{i<j}\frac{e^2}{r_{ij}} + \frac{Z_a Z_b e^2}{R} \tag{1.3}$$

where $r_{ai}$ and $r_{bi}$ are the distances between electron $i$ and nuclei $a$ and $b$ , having the charges $Z_a$ and $Z_b$, respectively, and $R$ is the fixed interatomic distance.

The wave function $\Psi_m(r,R)$ describes the electronic motion at fixed nuclear coordinates or at an infinitely slow change of nuclear coordinates. So, the eigenvalues $E_m$ of Equation (1.2) depend upon the value of parameter $R$. The solution of Equation (1.2) at different values of $R$ allows to obtain the function $E_m(R)$. This function plays the role of the potential energy in the Schrödinger equation for the nuclear motion, which in the so-called Born–Oppenheimer approximation[5] is presented as:

$$[-\frac{\hbar^2}{2\mu}\nabla_R^2 + E_m(R)]\chi_{mv}(R) = E_{mv}\chi_{mv}(R) \qquad (1.4)$$

As was noted by Born[6], the adiabatic potential energy $E_m(R)$ can be corrected by adding the diagonal contribution of the electron-nuclear interaction. In this Born adiabatic approximation, the Schrödinger equation for the nuclear motion is written as:

$$[-\frac{\hbar^2}{2\mu}\nabla_R^2 + V_m(R)]\chi_{mv}(R) = E_{mv}\chi_{mv}(R) \qquad (1.5)$$

$$V_m(R) = E_m(R) - W_{mm}(R) \qquad (1.6)$$

Here the adiabatic potential energy $V_m(R)$ is referred to as the interatomic potential. The second term in Equation (1.6), $W_{mm}(R)$, is called adiabatic correction. Because of computational complexity, the adiabatic corrections are often neglected and the interatomic potential is approximated by the energy $E_m(R)$ found in solution of the Schrödinger equation for the electronic motion at different values of the interatomic distance $R$.

The potentials describing the interaction between two molecules may be, as in the two-atom case, represented by a potential curve depending only on one variable, if

5

the interaction is averaged over all molecular orientations in space. This potential, $V(R)$ where $R$ is the distance between the center of masses of the molecules, is named the intermolecular potential. When the interacting molecules are fixed in space, the intermolecular potential, $V(R, \vartheta)$, depends on a set of Euler's angles, $\vartheta$, which determine the mutual orientation of the molecules. In addition, the interaction depends on the electronic state of the interacting systems. Therefore, for any pair of molecules there is a set of intermolecular potentials, $V_m(R, \vartheta)$, where $m$ labels the quantum states of the systems.

# 1.2 General classification of intermolecular interactions

The classification of intermolecular interactions depends on the distance between interacting objects. Various types of intermolecular interactions are listed in Figure 1-1. They are classified according to the three ranges of intermolecular separation where they play dominating role in the description of the interaction of molecules. These three ranges are:

I.   A range of short distances at which the potential has a repulsive nature.

II.  A range of intermediate distances with the van der Waals minimum, which is a result of the balance of the repulsive and attractive forces.

III. A range of large distances at which the electronic exchange is negligible and the intermolecular forces are attractive.

**Range I**

In this region, the perturbation theory (PT) for calculating the intermolecular interactions cannot be applied. To some extent, the interacting atoms (molecules) lose

their individuality because of a large overlap of their electronic shells. The same variational methods, which are used for molecular calculations, can be applied to the calculation of the total energy of interacting system, which can be considered as a 'supermolecule'. The interaction energy is found as a difference:

$$E_{int} = E_{tot} - \sum_{a=1}^{n} E_a \tag{1.7}$$

where $E_a$ is the energy of isolated subsystems (molecules or atoms) that have to be calculated at the same approximation as a whole system. In this region one can separate only two types of interaction energies: the Coulomb energy and the exchange energy. If zero is put for all integrals containing the exchange or overlap of electron densities, $E_{Coul}$ can be calculated. Then, the exchange energy is defined as the difference:

$$E_{exch} = E_{int} - E_{Coul} \tag{1.8}$$

**Range II**

Both repulsive and attractive forces exist in this region. This causes the minimum of intermolecular potential energy and provides the ground for the stability of the system. The magnitude of the interaction energy is much smaller than the self–energy of the interacting molecules and the perturbation theory can be applied although the electron exchange, appearing as a consequence of the antisymmetry of the total wave function, is still large. The standard perturbation Rayleigh–Schrödinger and Brillouin–Wigner theories are developed for the zero–order wave function taken as a simple product of the wave functions of the interacting molecules. In the intermediate distance region, the antisymmetric zero–order wave function has to be dealt with. This leads to essential modifications of the standard perturbation schemes. The approaches developed were named Symmetry Adapted Perturbation Theories (SAPT)[7].

The exchange energy is separated only in the first order of SAPT. The decomposition of $E_{int}$ into a perturbation series can be written as:

$$E_{int} = \varepsilon_{el}^{(1)} + \varepsilon_{exch}^{(1)} + \sum_{n=2}^{\infty} \varepsilon_{pol.exch}^{(n)} \qquad (1.9)$$

Where $\varepsilon_{el}^{(1)}$ is the classical electrostatic interaction energy between two (or more) systems of charges. The only difference to the classical expression is in the charge distribution: instead of point charges, $\varepsilon_{el}^{(1)}$ contains distributed electron densities. $\varepsilon_{exch}^{(1)}$ is the exchange energy in the first order of PT. Its origin is based on the Pauli principle demanding the antisymmetrization of many–electron wave functions. Because of antisymmetrization, electrons have the probability of being located on each interacting molecules. This is a specific quantum–mechanical effect. In the second and higher orders of SAPT, the exchange effects cannot be separated from the polarization energy.

**Range III**

In this region, exchange effects can be neglected. Usually, it is valid for $R \geq 15 a_0$. The different types of intermolecular interactions are classified in the frame of the standard Rayleigh–Schrödinger perturbation theory. In the second order of PT, the polarization energy splits on the induction $\varepsilon_{ind}^{(2)}$ and dispersion $\varepsilon_{disp}^{(2)}$ energies. In higher orders, some mixed term $\varepsilon_{ind,disp}$, which results from coupling of the induction and dispersion energies, appears. In the third order of PT, $\varepsilon_{ind,disp}^{(3)}$ was analyzed by Jeziorski et al.[7a]. At distances where the exchange effects are negligible, the PT expansion can be written:

$$E_{int} = \varepsilon_{el}^{(1)} + \sum_{n=2}^{\infty} [\varepsilon_{ind}^{(n)} + \varepsilon_{disp}^{(n)}] + \sum_{n=3}^{\infty} \varepsilon_{ind,disp}^{(n)} \qquad (1.10)$$

8

At these distances the multipole expansion of the electrostatic potential is valid. The interaction operator can be expanded in the multipole series. As a result, $\varepsilon_{el}^{(1)}$ represents a direct multipole–multipole electrostatic energy. At sufficiently large distances, only the first term in the multipole expansion is sufficient to describe the interaction. For polar molecules, it is the dipole term.

The physical sense of the induction energy $\varepsilon_{ind}^{(2)}$ is the same as in classical physics: the induction of the dipole electric moment in one molecule by the permanent quadrupole moment of the other molecule. The dispersion energy is a pure quantum-mechanical phenomenon. It originates in the quantum-mechanical fluctuations of electronic density. The instant redistribution of electron density leads to the nonzero mean dipole moment even in the cases when the permanent dipole moment is equal to zero (nonpolar molecules, noble gas atoms). This instant dipole moment induces a dipole and higher moments in the other molecule. In the second order of PT, the multipole expansion of the dispersion energy $\varepsilon_{disp}^{(2)}$ is usually written in the following form:

$$\varepsilon_{ind}^{(2)} = -\sum_{n=6}^{\infty} \frac{C_n}{R^n} \tag{1.11}$$

The coefficients $C_n$ are named as dispersion coefficients. In higher orders of PT, the interpretation of different terms in Equation (1.11) becomes more complex.

Magnetic interactions exist at all distances, but in ranges I and II they are often negligible in comparison with larger electrostatic interactions. On the other hand, magnetic interactions are precisely detected by measurements of the energy level splitting in a magnetic field. In some special cases, the magnetic interactions become the largest at sufficiently large distances. This is the case for oriented nonpolar

molecules possessing the quadrupole moment as the first nonvanishing multipole moment and having, in the ground state, the total electronic spin $S \neq 0$. The first term in the multipole expansion of different electrostatic intermolecular interactions for these molecules has the following distance dependence:

$\sim \dfrac{1}{R^5}$ in the electrostatic energy $\varepsilon_{el}^{(1)}$ (the quadrupole–quadrupole interaction).

$\sim \dfrac{1}{R^6}$ in the dispersion energy $\varepsilon_{disp}^{(2)}$ (the dipole–induced dipole interaction).

$\sim \dfrac{1}{R^8}$ in the induction energy $\varepsilon_{ind}^{(2)}$ (the quadrupole–induced dipole interaction).

The magnetic spin–spin interaction is relativistic one and has a dipole-dipole distance behavior. It is proportional to $\alpha^2 / R^3$, where $\alpha = 1/137$ is the fine structure constant. Although $\alpha^2$ is small, it is evident that the magnetic interactions become predominant with increasing intermolecular distance due to their asymptotics. The described situation is realized in example for oxygen molecules adsorbed on some surface. The oxygen molecule, as all homonuclear diatomic molecules, has no dipole moment and its ground state is triplet (total electronic spin $S = 1$[8]).

At distances at which the propagation time of the interaction, $R/c$, is of the same order as the average time of the electronic transitions, which is proportional to $\hbar / I_1$ (where $I_1$ is the first ionization potential), that is, $R \sim \hbar c / I$, the retardation effect should be taken into account. Usually, this effect becomes appreciable at $R \sim 500$ bohrs and this is not consider in the Thesis.
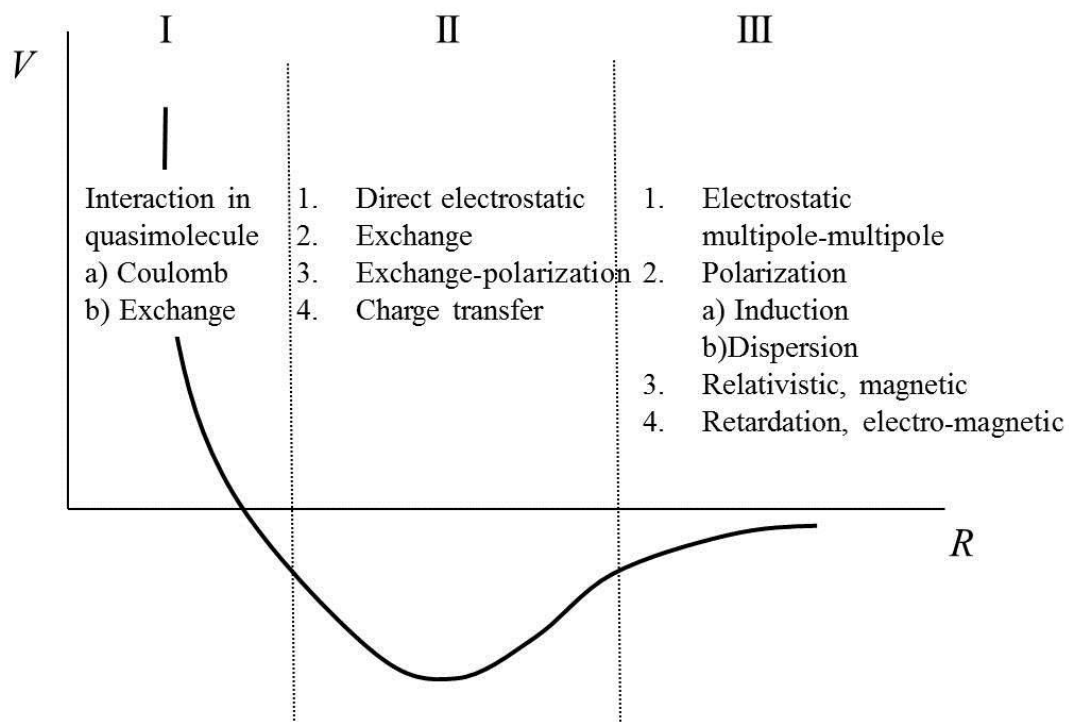
**Figure 1-1** Classification of intermolecular interactions.[9]
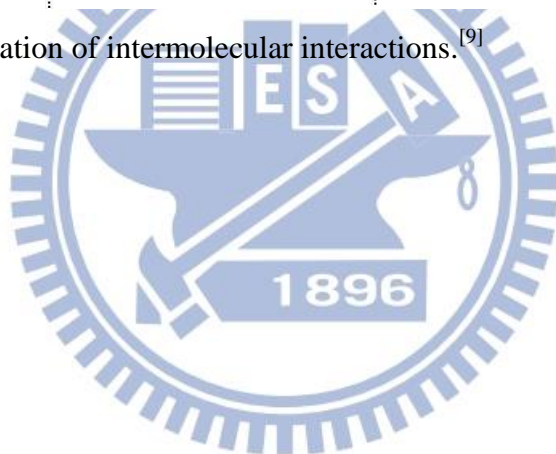
# Chapter 2

# Theory and method

## 2.1 Coupled–cluster Theory

Coupled–cluster (CC) theory is one of several post–Hartree–Fock ab initio quantum chemistry methods. It is nowadays one of the most popular methods in quantum chemistry that treats electronic correlation.

Coupled–cluster theory constructs solutions to the time–independent Schrödinger equation

$$\hat{H}|\Psi\rangle = E|\Psi\rangle$$

where $\hat{H}$ is the Hamiltonian of the system and the CC wavefunction $|\Psi\rangle$ is written in an exponential form

$$|\Psi\rangle = \exp(\hat{T})|\Phi_0\rangle$$

where $|\Phi_0\rangle$ is a Slater determinant constructed from Hartree–Fock molecular orbitals and $\hat{T}$ is the cluster operator.

The cluster operator is expressed as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ...$$

where $\hat{T}_1$ is the operator of all single excitations, $\hat{T}_2$ is the operator of all double excitations and so on. In the formalism of second quantization these excitation operator are written as

$$\hat{T}_1 = \sum_i \sum_a C_a^i \hat{a}^a \hat{a}_i,$$

$$\hat{T}_2 = \frac{1}{4} \sum_{i,j} \sum_{a,b} C_{ab}^{ij} \hat{a}^a \hat{a}^b \hat{a}_j \hat{a}_i,$$

and for the general $n$–fold cluster operator

$$\hat{T}_n = \frac{1}{(n!)^2} \sum_{i_1,i_2,...,i_n} \sum_{a_1,a_2,...,a_n} C_{a_1,a_2,...,a_n}^{i_1,i_2,...,i_n} \hat{a}^{a_1} \hat{a}^{a_2} \dots \hat{a}^{a_n} \hat{a}_{i_n} \dots \hat{a}_{i_2} \hat{a}_{i_1}$$

In the above formula, $\hat{a}^a$ and $\hat{a}_i$ are denoted as creation and annihilation operators, $i, j$ denote occupied orbitals and $a$, $b$ denote unoccupied orbitals. When the cluster operator is truncated at the two–electron excitations, i.e., when $\hat{T} = \hat{T}_1 + \hat{T}_2$, people usually denote the method as CCSD method. Continue to CCSD(T), CCSDT, etc.

## 2.2 Explicitly correlated coupled–cluster methods

The CCSD(T) method is known to have an excellent accuracy for many applications in quantum chemistry. Unfortunately, the steep $O(N^7)$ cost scaling of CCSD(T), where $N$ is a measure of the molecular size, combined with its strong basis set dependence, limits its applicability for obtaining highly accurate results to small molecules. The slow convergence of the electron correlation energy with respect to the basis set size can be improved by including terms into the wave function that depend explicitly on the interelectronic distances. In 2007, Werner *et al.* proposed[10] a simple and efficient CCSD(T)-F12$x$ approximations ($x$=a, b) that yields highly accurate results. Recently, the interaction energies for weakly bound dimer computed by several variants of the explicitly correlated CCSD(T)-F12 method has been reported[11]. It is shown that the F12 approach significantly speed up the convergence of the CCSD(T)/aug-cc-pV$X$Z interaction energies with the basis set cardinal number $X$.

In the following, the indices $i, j,..., p$ denote occupied orbitals, $a, b,..., d$

denote virtual orbitals and $\alpha, ..., \gamma$ are the (orthonormal) orbitals of a formally complete virtual space. This can be partitioned into the virtual orbital basis set $\{a, b, ...\}$ and the complementary auxiliary basis set (CABS) $\{x, y, ...\}$, so that $|\alpha\rangle\langle\alpha|$ can be approximated by $|a\rangle\langle a| + |x\rangle\langle x|$. The one-electron operators $\hat{o}_n = |i_n\rangle\langle i_n|$, $\hat{v}_n = |a_n\rangle\langle a_n|$ project onto the occupied and virtual orbital subspaces, and their subscripts refer to the electron coordinates on which they act.

The wave function employed in the full CCSD-F12 approach has the form $|\Psi\rangle_{\text{CCSD-F12}} = \exp(\hat{T}_1 + \hat{T}_2)|\Phi\rangle$, where $\Phi$ is the Hartree-Fock reference function. The single and double excitation cluster operators $\hat{T}_1$ and $\hat{T}_2$ are defined as

$$\hat{T}_1 = t_a^i \hat{E}_i^a,$$

$$\hat{T}_2 = \frac{1}{2} T_{ab}^{ij} \hat{E}_{ij}^{ab} + \frac{1}{2} \tau_{\alpha\beta}^{ij} \hat{E}_{ij}^{\alpha\beta},$$

where $\hat{E}_i^\alpha$ and $\hat{E}_{ij}^{\alpha\beta} = \hat{E}_i^\alpha \hat{E}_j^\beta$ are the usual spin-free one- and two-electron excitation operators. We use different notation here from previous section. The excitation into the standard virtual orbitals $a$, $b$ in $\hat{T}_1$ and the first part of $\hat{T}_2$ are the same as those used in the conventional CCSD theory. The additional $\tau_{\alpha\beta}^{ij}$ amplitudes are given by

$$\tau_{\alpha\beta}^{ij} = T_{mn}^{ij} F_{\alpha\beta}^{mn},$$

$$F_{\alpha\beta}^{mn} = \langle mn | F_{12} \hat{Q}_{12} | \alpha\beta \rangle,$$

$$\hat{Q}_{12} = (1 - \hat{o}_1)(1 - \hat{o}_2)(1 - \hat{v}_1 \hat{v}_2),$$

where $F_{12} \equiv F(r_{12})$ is a short-range correlation factor and $T_{mn}^{ij}$ are the actual amplitudes used in the F12 treatment. The projector $\hat{Q}_{12}$ is necessary to make the F12

configurations, $\left|\Phi_{ij}^{mn}\right\rangle = F_{\alpha\beta}^{mn} E_{ij}^{\alpha\beta}\left|\Phi\right\rangle$ , orthogonal to the configurations in the molecular orbital space. Thus, $F_{\alpha\beta}^{mn} = 0$, and therefore $\tau_{rs}^{ij} = 0$.

In practice people use an exponential correlation factor fitted to a set of Gaussian geminals, $F(r_{12}) = -\dfrac{1}{\gamma}\exp(-\gamma r_{12}) \approx \sum_i c_i \exp(-\alpha_i r_{12}^2)$ .

The above ansatz is orbital invariant. However, it suffers from germinal basis set superposition errors, and for larger molecules also numerical instabilities occur. In order to avoid these problems, people use the "diagonal" ansatz. The short-range asymptotic behavior of the correlation cusp of the exact wave function can be used to great advantage in the F12 theory. As pointed out by Tenno[12], a wave function $T_{mn}^{ij}\left|\Phi_{ij}^{mn}\right\rangle$ with amplitudes $T_{mn}^{ij}$ fixed to $T_{ii}^{ii} = t_s$, $T_{ij}^{ij} = \dfrac{1}{2}(t_s + t_t)$, $T_{ji}^{ij} = \dfrac{1}{2}(t_s - t_t)$, where $i \neq j$, $t_s = \dfrac{1}{2}$, $t_t = \dfrac{1}{4}$, and the remaining amplitudes are set to zero, fulfills these asymptotic conditions.

To compensate for the lack of an F12 triples contribution, a scaling approach has been suggested[13]

$$E_{(T)}^* = E_{(T)} \frac{E_{\text{MP2-F12}}^{\text{corr}}}{E_{\text{MP2}}^{\text{corr}}},$$

where $E_X^{\text{corr}}$ denotes correlation energy at the X level of theory, i.e., $E_X - E_{\text{SCF}}$.

## 2.3 Complete basis set limit extrapolation

The complete basis set (CBS) limit does not correspond to any existing basis set though it is often written in such a form, e.g. CCSD(T)/CBS. Instead, the CBS limit is obtained by extrapolating the results which come from large basis set calculations. For many properties the CCSD(T)/CBS value can be regarded as a numerically exact for all practical purposes, i.e. it is unlikely that any higher level of theory predicts

significantly better results.

The extrapolation is based on a minimum of two separate calculations with increasingly larger basis sets. CBS limit extrapolation works only with basis sets designed specifically for the task, such as the correlation– or polarization–consistent basis sets, e.g. cc–pV$XZ$ or pc–$n$.

The procedure is as follows: a given property $E$ of interest (e.g. a relative energy, a frequency, or a bond length) is computed at a given level of theory (e.g. B3LYP) using three basis sets (e.g. cc–VDZ, cc–VTZ, and cc–VQZ). These data points are then fit to an equation, the two most popular equations are given here

$$E(x) = E_{CBS} + Ae^{-(C \cdot X)} \tag{2.1}$$

$$E(x) = E_{CBS} + AX^{-3} \tag{2.2}$$

Here, $E_{CBS}$ is the CBS limit we want to determine and $X$ is 2 for cc–pVDZ, 3 for cc–pVTZ, and so on. $X$ is also often denoted as $L_{max}$ (or "cardinal number"), which is the highest angular momentum included in the basis set. For cc–pVDZ this means d orbitals, which have an angular momentum of 2, so $X$ and $L_{max}$ are really the same.

Equation (2.1) contains three parameters ($E_{CBS}$, $A$, and $C$) so a minimum of three different basis sets are needed to determine them. While Equation (2.2) only has two parameters, a minimum of two data points are needed.

For some properties and correlation methods the use of the low–zeta basis sets does not provide high accuracy, so calculations with even pentuple–zeta basis sets are sometimes necessary. CBS limit extrapolation is computationally very demanding and is typically done on relatively small systems to provide benchmark values.

## 2.4 Basis set superposition error

The intermolecular interaction energy is defined as a difference of the total energy of the 'supermolecule' and the energies of constituent molecules. For a two-molecule system, we have

$$E_{\text{int}} = E(AB) - [E(A) + E(B)] \qquad (2.3)$$

The supermolecular approach is valid for any intermolecular separations and this is its advantage in comparison with the perturbation approaches. The determination of two large quantities, imposes strict requirements on the accuracy of the energies. The calculation of the interaction energy in the supermolecular approach is a subject to one serious defect connected with a basis set inconsistency leading to an artificial enhancement of the intermolecular interaction energy. Consider this problem in the case of a dimer AB. In the standard approach, it is natural to define:

$$E_{int}^{St}(R) = E^{AB}(R) - [E^A(\{A\}, \infty) + E^B(\{B\}, \infty)] \qquad (2.4)$$

where $R$ is the separation between monomers in the dimer AB and $E^A(\{A\}, \infty)$ ($E^B(\{B\}, \infty)$) denotes the monomer A (B) energy calculated at infinite separation using only the monomer basis set $\{A\}$ ($\{B\}$), respectively. Note that in practice the monomer basis is never complete. The unified dimer basis $\{AB\}$ is larger than that of monomer A from the point of view of monomer A. This cause an artificial energy stabilization deepening the potential energy curve of the dimer.

The solution for this defect was proposed by Jansen and Ros[14] for a particular reaction and, as a general method, by Boys and Bernardi[15]. They suggested that the supermolecular basis is used for all terms in Equation (2.4). Their approach was called function counterpoise[15]; it is usually designated by an abbreviation CP:

$$E_{int}^{CP}(R) = E^{AB}(R) - [E^A(\{AB\}, R) + E^B(\{AB\}, R)] \qquad (2.5)$$

where $E^A(\{AB\}, R)$ is the energy of the monomer A calculated using the dimer basis set $\{AB\}$. It means that A is calculated using not only its own basis set $\{A\}$ but also a set of ghost orbitals $\{B\}$ centered at the distance $R$ from A. In practice, this calculation is formally identical to the dimer calculation except that the nuclear charges are put at zero for all nuclei belonging to B and an appropriate number of electrons is subtracted from the calculation.

Thus, the CP procedure adds an 'extra' stability to monomers that makes their calculation consistent with the dimer calculation. The error in Equation (2.4) caused by the basis inconsistency is named the basis set superposition error (BSSE)[16]. It is defined as the difference

$$BSSE(R) = E_{int}^{CP} - E_{int}^{St} = [E^A(\{A\}, \infty) + E^B(\{B\}, \infty)] - [E^A(\{AB\}, R) + E^B(\{AB\}, R)]$$

## 2.5 Midbond function technique

This technique was proposed by Tao and Pan[17]. It enables the calculation of accurate interaction energies and equilibrium distances for dimers in a substantially smaller basis set by adding some basis functions centered between the subsystems. The so-called midbond functions lead to faster convergence than high angular momentum functions centered on the nuclei.

As noted by Tao and Pan[17a], the major contribution to the correlation interaction energy comes from double excitations. Two types of double excitations may be identified. First, both excited electrons from one subsystem. Second, only one electron originates from each system. The first produces the intrasystem correlation interaction energy, which is either repulsive or attractive and relatively small, and the

second produces the intersystem correlation interaction energy, which is always attractive and dominant.

The intersystem correlation energy may be analogous to the regular bonding energy. Consider two isolated atoms or molecules A and B, each with an excited electron in its own virtual orbital space $\phi_0$. When the two systems approach each other to form the complex AB, the interaction of $\phi_0$ between the two systems results in the formation of the "bonding" molecular orbital, $\psi$, which is to be occupied by the two excited electrons. This produces the energy lowering, $\Delta\varepsilon$, which is a contribution to the intersystem correlation energy and is quite similar to the regular bonding energy. The resulting molecular orbital $\psi$ should have the following characteristics: it is located mainly in the region between A and B, and it is highly diffuse. In a normal calculation, very diffuse and highly polarized functions were used to provide the adequate description of the orbital $\psi$. This is one of the key reasons for using large basis sets for intermolecular energy calculations. A possible and straightforward way is to place some basis functions at the midway of the van der Waals bond between A and B. These midway functions are referred to as midbond functions. Such midbond functions are more efficient for the description of the orbital $\psi$ than the functions centered on the nuclei of A and B. The above description is illustrated in Figure 2-1.

Some precautions should be considered when using midbond functions. First, the nucleus-centered basis functions describing the core and valence electrons should be approximately saturated before introducing the midbond functions. This means that the BSSE of the basis set at the Hartree–Fock level should be sufficiently small. Otherwise, severe distortion of electron distribution may result, which in turn causes

incorrigible higher-order BSSE. Second, some of the nucleus-centered polarization functions are still needed for the description of intrasystem correlation effects since the midbond functions are usually limited to provide the description only for the intersystem correlation effects.
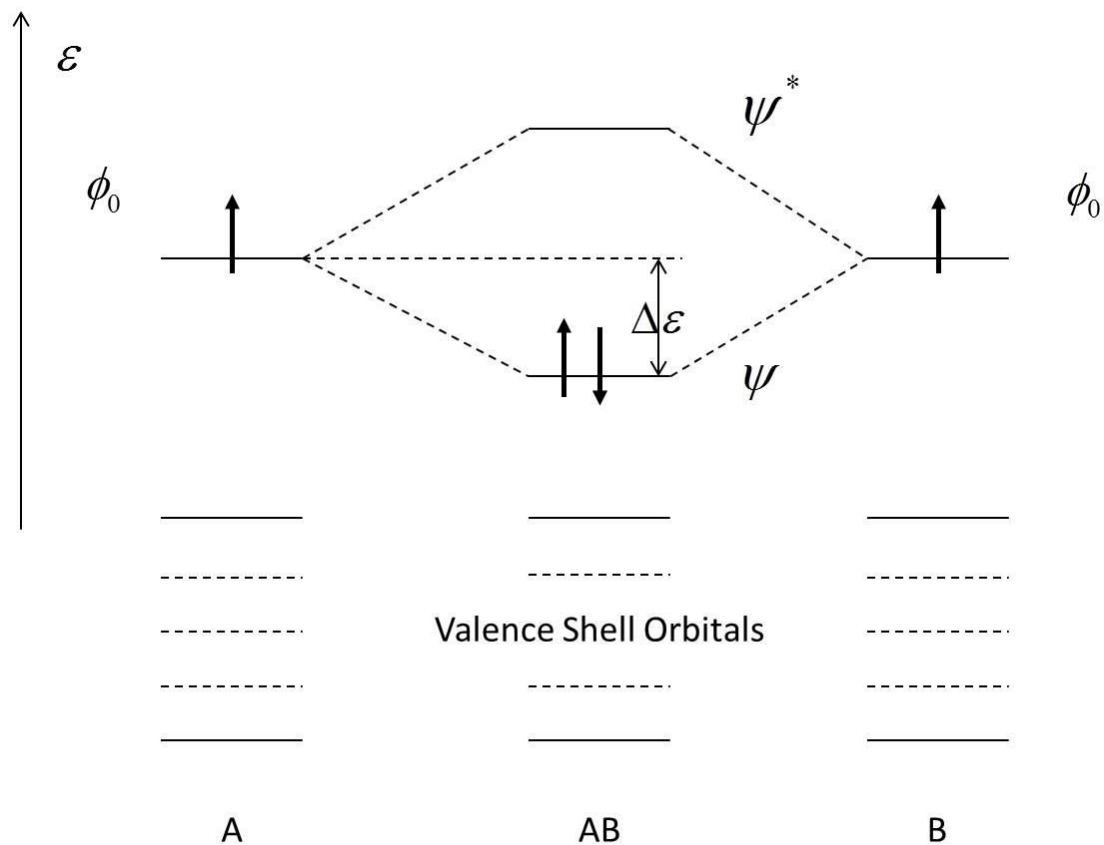


**Figure 2-1** Graphic representation of the molecular-orbital interactions in the doubly excited configurations of the complex AB[17a].

## 2.6 Vibrationally averaged interaction energy surfaces

One of the challenges of the theory of intermolecular interactions is to go beyond the rigid monomers approximation. The time required for constructing full-dimensional interaction energy potential is a product of two factors: the time required to get interaction energy for one spatial configuration and the number of points of the grid. The total time of calculations would be prohibitive if the interaction energy is calculated by very accurate method. In order to take into account the dependence of interaction energy on intramolecular coordinates efficiently, a method called vibrationally averaged interaction energy for the generation of a reliable interaction potential energy surfaces has been proposed[1] and tested[1-3].

A characteristic feature of any van der Waals complex is that the constituent fragments of the complex preserve their identity. Thus one can distinguish among the coordinates describing the geometry of the whole system, the set $\mathbf{X}$ of N intermolecular coordinates $X_i$ and the set $\mathbf{x}$ of $n$ intramolecular coordinates $x_i$. The approximation $V_{AP}(\mathbf{X},\mathbf{x})$ to the exact interaction energy surface $V(\mathbf{X},\mathbf{x})$ is represented as the truncated Taylor expansion around some reference intramolecular configuration $\mathbf{x_c}$,

$$V_{AP}(\mathbf{X},\mathbf{x}) \approx V(\mathbf{X},\mathbf{x_c}) + \left(\frac{\partial V}{\partial q_i}\right)_{\mathbf{q}=0} q_i + \frac{1}{2}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_{\mathbf{q}=0} q_i q_j, \qquad (2.6)$$

where the set $\mathbf{q}$ consists of the components $q_i = x_i - x_{ci}$, and $x_{ci}$ is the $i$ th element of $\mathbf{x_c}$. The Einstein summation convention is used for Latin indices $i, j, \ldots$.

The partial derivatives computed for $\mathbf{q} = 0$ in the equation above depend on the intermolecular coordinates $\mathbf{X}$ and define the coefficients of the expansion. Thus, as one has the knowledge of the functional form of these partial derivatives, one has the full information needed to calculate the values of $V_{AP}(\mathbf{X},\mathbf{x})$. In many cases, the dependence of $V(\mathbf{X},\mathbf{x})$ on the intramolecular coordinates $\mathbf{x}$ is weak. So, the Taylor expansion can be truncated to only a few leading terms.

The $V(\mathbf{X},\mathbf{x})$ interaction potential energy surface is defined as $V(\mathbf{X},\mathbf{x}) = E_{AB}(\mathbf{X},\mathbf{x}) - E_A(\mathbf{x_A}) - E_B(\mathbf{x_B})$, where $E_{AB}$ is the total energy of the dimer, $E_A$ ($E_B$) is the total energy of monomer A (B), and $\mathbf{x_A}$ ($\mathbf{x_B}$) is the set of internal coordinates of the monomer A (B).

The dimensionality of $V_{AP}(\mathbf{X},\mathbf{x})$ surface can be further reduced by averaging it over definite vibrational states $v$ of the subsystems A and/or B, which yields the following formula

$$\left\langle V_{AP}(\mathbf{X}) \right\rangle_v \approx V(\mathbf{X},\mathbf{x_c}) + \left(\frac{\partial V}{\partial q_i}\right)_{\mathbf{q}=0} \left\langle q_i \right\rangle_v + \frac{1}{2}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_{\mathbf{q}=0} \left\langle q_i q_j \right\rangle_v . \qquad (2.7)$$

This equation shows that to calculate the vibrationally averaged $\left\langle V_{AP}(\mathbf{X}) \right\rangle_v$ potential, one needs to know only a limited number of the potential derivatives and the values of $\left\langle q_i \right\rangle_v$ and $\left\langle q_i q_j \right\rangle_v$. $\left\langle q_i \right\rangle_v$ and $\left\langle q_i q_j \right\rangle_v$ can be readily obtained from the theoretical or empirical data for monomers.

In this Thesis, the set $\mathbf{X} = \{R, \theta_1, \theta_2, \phi\}$ and the set $\mathbf{x} = \{r, s\}$ (see Ch.3.1). The interaction energy of the H$_2$–HCl complex, $V(\mathbf{X},\mathbf{x}) = V(R, \theta_1, \theta_2, \phi, r, s)$, can be written as

$$
\begin{aligned}
V(\mathbf{X},r,s) \approx\; & f_{00}(\mathbf{X}) \\
& + f_{10}(\mathbf{X})(r-r_c) + f_{01}(\mathbf{X})(s-s_c) + f_{11}(\mathbf{X})(r-r_c)(s-s_c) \\
& + \frac{1}{2} f_{20}(\mathbf{X})(r-r_c)^2 + \frac{1}{2} f_{02}(\mathbf{X})(s-s_c)^2.
\end{aligned}
\tag{2.8}
$$

The derivatives $f_{ij}$ are calculated numerically with steps $h_r$ and $h_s$ from the following formulas:

$$
f_{00}(\mathbf{X}) = V(\mathbf{X},r_c,s_c),
\tag{2.9}
$$

$$
f_{10}(\mathbf{X}) = \left( \frac{\partial V}{\partial r} \right)_{r=r_c} \approx \frac{V(\mathbf{X},r_c+h_r,s_c) - V(\mathbf{X},r_c-h_r,s_c)}{2h_r},
\tag{2.10}
$$

$$
f_{01}(\mathbf{X}) = \left( \frac{\partial V}{\partial s} \right)_{s=s_c} \approx \frac{V(\mathbf{X},r_c,s_c+h_s) - V(\mathbf{X},r_c,s_c-h_s)}{2h_s},
\tag{2.11}
$$

$$
f_{20}(\mathbf{X}) = \left( \frac{\partial^2 V}{\partial r^2} \right)_{r=r_c} \approx \frac{V(\mathbf{X},r_c+h_r,s_c) - 2V(\mathbf{X},r_c,s_c) + V(\mathbf{X},r_c-h_r,s_c)}{h_r^2},
\tag{2.12}
$$

$$
f_{02}(\mathbf{X}) = \left( \frac{\partial^2 V}{\partial s^2} \right)_{s=s_c} \approx \frac{V(\mathbf{X},r_c,s_c+h_s) - 2V(\mathbf{X},r_c,s_c) + V(\mathbf{X},r_c,s_c-h_s)}{h_s^2},
\tag{2.13}
$$

$$
\begin{aligned}
f_{11}(\mathbf{X}) = & \left( \frac{\partial^2 V}{\partial s \partial r} \right)_{r=r_c,s=s_c} \\
& \approx \frac{V(\mathbf{X},r_c+h_r,s_c+h_s) - V(\mathbf{X},r_c+h_r,s_c) - V(\mathbf{X},r_c,s_c+h_s) + V(\mathbf{X},r_c,s_c)}{h_r h_s}.
\end{aligned}
\tag{2.14}
$$

Later, the full-dimensional potential energy surface will be vibrationally averaged over the vibrational wave functions of a molecule A, $\chi_v$, and B, $\chi'_{v'}$,

$$
\begin{aligned}
\langle V \rangle_{vv'}(\mathbf{X}) = & \left\langle \chi_v(r)\chi'_{v'}(s) \middle| V(\mathbf{X},r,s) \middle| \chi_v(r)\chi'_{v'}(s) \right\rangle \\
\approx\; & f_{00}(\mathbf{X}) + f_{10}(\mathbf{X})(\langle r \rangle_v - r_c) + f_{01}(\mathbf{X})(\langle s \rangle_{v'} - s_c) \\
& + f_{11}(\mathbf{X})(\langle r \rangle_v - r_c)(\langle s \rangle_{v'} - s_c) \\
& + \frac{1}{2} f_{20}(\mathbf{X})(\langle r^2 \rangle_v - 2r_c\langle r \rangle_v + r_c^2) \\
& + \frac{1}{2} f_{02}(\mathbf{X})(\langle s^2 \rangle_{v'} - 2s_c\langle s \rangle_{v'} + s_c^2),
\end{aligned}
\tag{2.15}
$$

where $\qquad \left\langle r^n \right\rangle_v = \left\langle \chi_v(r) \middle| r^n \middle| \chi_v(r) \right\rangle \qquad$ and $\qquad \left\langle s^n \right\rangle_{v'} = \left\langle \chi'_{v'}(s) \middle| s^n \middle| \chi'_{v'}(s) \right\rangle \qquad$ denote

vibrationally averaged values of powers of intramonomer distances. The result

$\left\langle V \right\rangle_{vv'} (\mathbf{X})$ is a four-dimensional functional of the following variables $\left\{ R, \theta_1, \theta_2, \phi \right\}$.

The ground vibrational states of both HCl and $H_2$ molecules were used ($v = v' = 0$) in

the Thesis. Thus, the resulting surface will be denoted as $\left\langle V \right\rangle_{00}$.
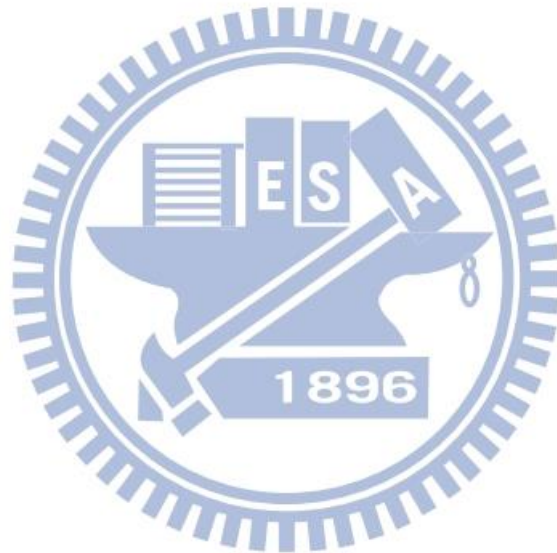
# Chapter 3

# Computational details

All calculations were performed with MOLPRO (version 2012) package[18] and MRCC program[19]. The values of the total interaction energy, $E_{int}$, of the $H_2$–HCl complex have been obtained by the supermolecular approach, i.e. $E_{int} = E^{H_2\text{-HCl}} - E^{H_2} - E^{HCl}$. All interaction energies were corrected for basis set superposition error using the counterpoise correction of Boys and Bernardi[15, 20]. The relativistic corrections were performed by using the Douglas-Kroll-Hess Hamiltonian[21]. Unless otherwise specified, the core electrons were not correlated. The basis sets used in this Thesis are Dunning's correlation consistent basis sets aug-cc-pV$X$Z[22] (abbreviated as AV$X$Z), cc-pV$X$Z (V$X$Z), and the tight-d augmented sets aug-cc-pV($X$+d)Z[23] (AV($X$+d)Z). The correlation consistent basis set for core correlation aug-cc-pCV$X$Z[24] (ACV$X$Z) and the weighted sets aug-cc-pwCV$X$Z[24] (AWCV$X$Z), are also used for core electrons correlation calculations.

For CCSD(T)–F12b[10, 13] calculations, AV$X$Z/MP2FIT auxiliary bases[25], with $X$ the same as for the orbital set, were used to fit the MP2-F12 pair functions. For the resolution of the identity (RI) approximation to many-electron, as well as for the density fitting of the Fock matrix, the AV$X$Z/jkfit auxiliary basis[26] were used. The default diagonal fixed-amplitude ansatz 3C(FIX)[27] was used. Exponent for Slater-type frozen geminal, or parameter for weight function in other frozen geminal models, $\beta$, was set to be $1.0\ a_0^{-1}$. The interaction energies obtained both with and without the scaling of the triples contribution will be presented. In the former case, the

scaling factor determined for the dimer was also used for both monomers to maintain size consistency[28].

The finite-basis results were extrapolated to the complete basis set (CBS) limit using the standard $X^{-3}$ formula[29] if the extrapolation technique is used. Specifically, the interaction energy extrapolated from basis sets $(X-1)Z$ and $XZ$ (this extrapolation will be denoted as CBS$(X-1, X)$) is a sum of the self-consistent field (SCF) contribution $E_{int}^{SCF}$ computed in the larger $XZ$ set and the correlation contribution $E_{int}^{corr}(CBS)$ obtained from the computed correlation energies $E_{int}^{corr}((X-1)Z)$ and $E_{int}^{corr}(XZ)$ as

$$E_{int}^{corr}(CBS)=E_{int}^{corr}(XZ)+\frac{(1-\frac{1}{X})^3}{1-(1-\frac{1}{X})^3}\cdot(E_{int}^{corr}(XZ)-E_{int}^{corr}((X-1)Z)).$$

While the $X^{-3}$ formula has been extensively used in conventional CCSD(T) calculations, it might not be the best choice for the CCSD(T)-F12b calculations. Theoretical considerations shows that $X^{-7}$ extrapolation scheme should be suitable here[30]. In my case, tests show that there is no systematic behavior from $X^{-3}$ to $X^{-7}$ extrapolation scheme. Therefore, the $X^{-3}$ formula still be used as an empirically justified way to improve basis set convergence[11].

The midbond functions were chosen as hydrogenic functions when they were used. In CCSD(T)–F12b calculations, the additional auxiliary basis functions for density-fitting and RI were chosen as described above. In other words, the midbond set varies with X in accordance with the atomic basis sets. The basis set symbol will be added "M", e.g. AVTZ+M, to indicate the midbond functions are included. In rare cases the midbond functions have been removed when convergence problem occured

in CCSD(T)–F12b calculations. This happened at short R distances (R=5.5, 6, 6.5) for the strongly repulsive part of the potential energy surface.

## 3.1 Coordinate system

The geometry of the $H_2$–HCl is described by six variables using Jacobi coordinates. $R$ denoting the distance between the center of the mass of $H_2$, CM($H_2$), and the center of the mass of HCl, CM(HCl). To define the angles $\theta_1$, $\theta_2$, and $\phi$, one can choose an axis, z, pointing from CM($H_2$) to CM(HCl). $\theta_1$ is the angle between z axis and the vector pointing from CM($H_2$) to the H atom on $H_2$ (H1). $\theta_2$ is the angle between z axis and the vector pointing from CM(HCl) to Cl. $\phi$ is the dihedral angle between two planes extending from z axis to H1 and to Cl. $r$ and $s$ are bond lengths of HCl and $H_2$, respectively. $R$, $\theta_1$, $\theta_2$, and $\phi$ are called intermolecular coordinates. $r$ and $s$ are called intramolecular coordinates. The midbond functions were chosen to be H atom functions which centered halfway between the projection of H atom from $H_2$ and the projection of H or Cl atom from HCl on the z axis. The system is plotted in Figure 3-1.
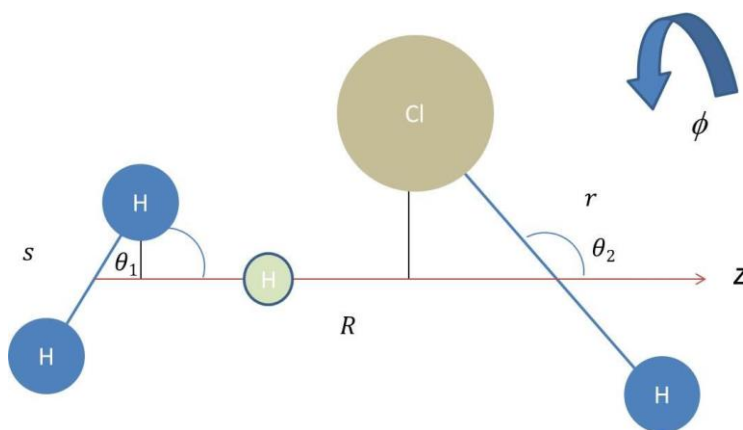


**Figure 3-1** Jacobi coordinates system of $H_2$–HCl.

## 3.2 Grid points

To obtain approximate full-dimensional potential energy surface for $H_2$–HCl, the leading $f_{00}$ term of expansion (Eqs. (2.9)) and the other $f_{ij}$ coefficients (Eqs. (2.10) – (2.14)) need to be calculated. To calculate $f_{00}$ term, one has to perform calculations for $(r_c, s_c)$, while to compute the $f_{ij}$ coefficients one has to use all six grid points in $(r, s)$, as shown in Eqs. (2.10) – (2.14). These six points are defined by the values of the steps $h_r = h_s = 0.025$ bohr. The Taylor expansions have been performed around the points $r_c = 2.434$ and $s_c = 1.474$ bohrs for HCl and $H_2$, respectively. $r_c$ is chosen close to vibrationaly averaged bond length for the vibrational ground state. $s_c$ is chosen to lie between $\langle s \rangle_0$ and $\langle s \rangle_1$. The value of $\langle r \rangle_0 = 2.4391$ bohr, $\langle r^2 \rangle_0 = 5.9703$ bohr were calculated by P. Jankowski[31] using analytical potential for HCl molecule[32] using modified version of the TRIATOM program[33]. $\langle s \rangle_0 = 1.44874$ bohrs and $\langle s^2 \rangle_0 = 2.12705$ bohrs², taken from Ref. [34].

The grid was chosen by combining the following values of $R$ (in bohr): 5.5, 6, 6.25, 6.5, 6.75, 7, 7.5, 8, 8.5, 9, 10, 11, 12, 15, 20 with the set of 33 unique combinations of the following angles: 0°, 45°, 90° and 135° for $\theta_1$; 0°, 45°, 90°, 135° and 180° for $\theta_2$; and 0°, 45° and 90° for $\phi$. For three intermolecular distance, 6.5, 7.5, and 9 bohrs, the angular grid was made twice as dense in each of the $\theta_1$, $\theta_2$ and $\phi$ coordinates, which resulted in 208 additional angular configurations to enhance the

fitting performance for the angular dependences of the potential energy surface near the region of its' minimum. Total of 1119 grid points were chosen in $X=\{R, \theta_1, \theta_2, \phi\}$ intermolecular coordinates.

Besides, in order to improve the performance of the fit, additional 15495 grid points were calculated for $(r_c, s_c)$ at CCSD(T)-F12b/AVTZ level. These grid was chosen by combining the following angles: $0°$, $22.5°$, $45°$, $67.5°$, $90°$, $112.5°$, $135°$, and $157.5°$ for $\theta_1$; $0°$, $5.625°$, $11.25°$, $16.875°$, …, $174.375°$ for $\theta_2$; and $0°$, $22.5°$, $45°$, $67.5°$ and $90°$ for $\phi$. The values of $R$ were the same as listed above. The details about the fitting procedure would be presented in next section.

# 3.3 Analytic fit

In order to improve the performance of the fit without a large numerical effort, the final potential $\langle V \rangle_{00}$ is not directly fitted from the most accurate 1119 grid points. The procedures were listed in the following and in Figure 3-2:

1. Computing all coefficients in Eqs. (2.9) – (2.14) of 1119 grid points, later did the average procedure as shown in Eqs. (2.15). This step generated a set of energies called $\langle E \rangle_{00}$.

2. Computing additional 15495 grid points for $(r_c, s_c)$ at CCSD(T)-F12b/AVTZ level. The set of energies is called $E_{F12b/TZ}$. Fit $E_{F12b/TZ}$ to generate the analytic representation of the potential energy surface called $V_{F12b/TZ}$.

3. Selecting the same 1119 grid points from 15495 grid points in $\{R, \theta_1, \theta_2, \phi\}$

coordinates. Calculating the difference $\Delta E = \langle E \rangle_{00} - E_{F12b/TZ}$. Fit $\Delta E$ to generate the analytic representation of the potential energy surface called $\Delta V$.

4. The resulting potential energy surface $\langle V \rangle_{00} = V_{F12b/TZ} + \Delta V$.

There were two analytic representations of the potential energy surface have been chosen:

$$V_{F12b/TZ}(R, \theta_1, \theta_2, \phi) = V_{F12b/TZ}(R; \theta_1, \theta_2, \phi) = \sum_{i=6,8,10,12,14,16,18,20,11,13,15} \frac{a_i(\theta_1, \theta_2, \phi)}{R^i}$$

$$
\begin{aligned}
a_i(\theta_1, \theta_2, \phi) &= a_i(\theta_1; \theta_2, \phi) \\
&= b_{i,0}(\theta_2, \phi) + b_{i,1}(\theta_2, \phi)\cos(2\theta_1) + b_{i,2}(\theta_2, \phi)\cos(4\theta_1) + b_{i,3}(\theta_2, \phi)\cos(6\theta_1) \\
&\quad + b_{i,-1}(\theta_2, \phi)\sin(2\theta_1) + b_{i,-2}(\theta_2, \phi)\sin(4\theta_1) + b_{i,-3}(\theta_2, \phi)\sin(6\theta_1) \\
&= \sum_{j=-3}^{3} b_{i,j} f_j(2|j|\theta_1)
\end{aligned}
$$

$$b_{i,j}(\theta_2, \phi) = b_{i,j}(\phi; \theta_2) = \sum_{k=0}^{2} c_{i,j,k}(\theta_2)\cos(k\phi)$$

$$c_{i,j,k}(\theta_2) = \sum_{n=0}^{6} d_{i,j,k,n}\cos(n\theta_2) + \sum_{n=1}^{6} d_{i,j,k,-n}\sin(n\theta_2) = \sum_{n=-6}^{6} d_{i,j,k,n} f_n(n\theta_2)$$

$$V_{F12b/TZ}(R, \theta_1, \theta_2, \phi) = \sum_{i,j,k,n} d_{i,j,k,n} \frac{f_j(2j\theta_1) f_n(n\theta_2)\cos(k\phi)}{R^i}$$

And the expansion for $\Delta V$

$$\Delta V(R, \theta_1, \theta_2, \phi) = \Delta V(R; \theta_1, \theta_2, \phi) = \sum_{i=6,8,10,12,14,16} \frac{a_i(\theta_1, \theta_2, \phi)}{R^i}$$

$$
\begin{aligned}
a_i(\theta_1, \theta_2, \phi) &= a_i(\theta_1; \theta_2, \phi) \\
&= b_{i,0}(\theta_2, \phi) + b_{i,1}(\theta_2, \phi)\cos(2\theta_1) + b_{i,2}(\theta_2, \phi)\cos(4\theta_1) \\
&\quad + b_{i,-1}(\theta_2, \phi)\sin(2\theta_1) + b_{i,-2}(\theta_2, \phi)\sin(4\theta_1) \\
&= \sum_{j=-2}^{2} b_{i,j} f_j(2|j|\theta_1)
\end{aligned}
$$

$$b_{i,j}(\theta_2, \phi) = b_{i,j}(\phi; \theta_2) = \sum_{k=0}^{2} c_{i,j,k}(\theta_2)\cos(k\phi)$$

$$
\begin{aligned}
c_{i,j,k}(\theta_2) &= d_{i,j,k,0} + d_{i,j,k,1}\cos(\theta_2) + d_{i,j,k,2}\cos(2\theta_2) + d_{i,j,k,-1}\sin(\theta_2) + d_{i,j,k,-2}\sin(2\theta_2) \\
&= \sum_{n=-2}^{2} d_{i,j,k,n} f_n(|n|\theta_2)
\end{aligned}
$$

30

$$\Delta V(R,\theta_1,\theta_2,\phi) = \sum_{i,j,k,n} d_{i,j,k,n} \frac{f_j(2j\theta_1)f_n(n\theta_2)\cos(k\phi)}{R^i}$$

A FORTRAN program computing the fitted surfaces was developed by Michał Slawik[35] and be used in the Thesis.
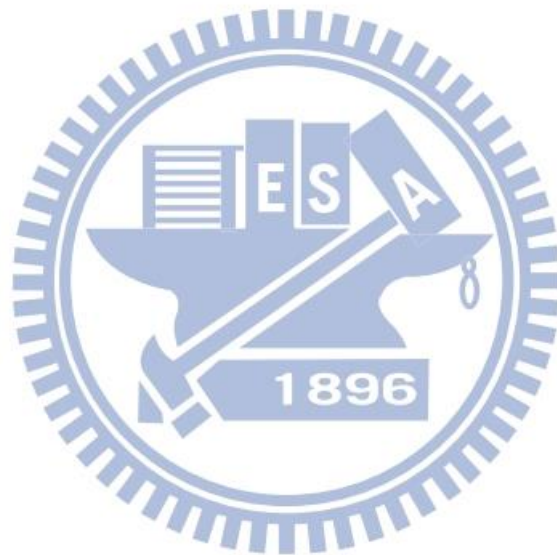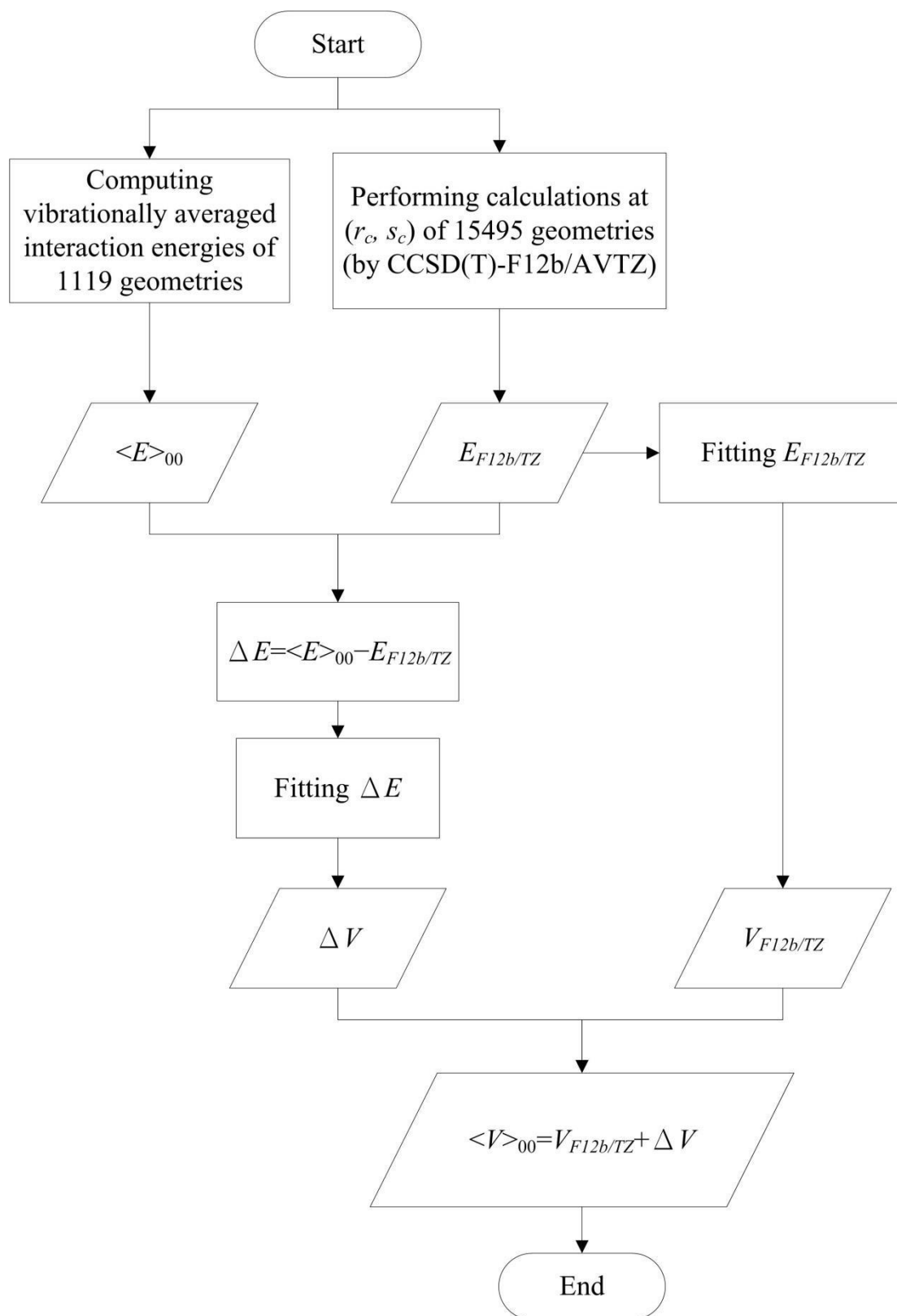
**Figure 3-2** The fitting procedure.

# Chapter 4

# Results and discussions

## 4.1 Benchmark calculations of interaction energies

In the following sections, the interaction energies at $\{R, \theta_1, \theta_2, \phi, r, s\} = \{6.5, 90, 0, 0, 2.409, 1.449\}$ will be discussed with respect to different levels of CC theory and different basis sets in order to find the most suitable theory level and basis set combinations to calculate the interaction energy of $H_2$–HCl complex. The main reason for choosing this geometry is that it is located on the grid points $\{R, \theta_1, \theta_2, \phi\}$ close to the interaction energy minimum. The interaction energies changed with varying basis set at this geometries is relatively large than at other geometries'.

## 4.1.1 CCSD(T) part

The CCSD(T) and CCSD(T)-F12b interaction energies computed in different basis sets with and without midbond functions, are presented in Table 4-1. Table 4-1 displays conventional CCSD(T) results for basis sets AV$X$Z, where $X$=D–6 and CCSD(T)-F12b results, with and without the scaling of triples, for basis sets AV$X$Z (X=D–5).

Results shown that for both the conventional CCSD(T) method and explicitly correlated extension of CCSD(T), obtained with addition of mindbond functions converge much faster than the results computed without midbond functions. The

CCSD(T)/CBS(5-6)+M result which derived from two largest AV5Z+M and AV6Z+M basis sets can be considered as the benchmark value and the difference between the value of CCSD(T)/CBS(5-6)+M and CCSD(T)/CBS(Q-5)+M can be considered as the error, then we get the $-206.080 \pm 0.4$ cm$^{-1}$ as the benchmark value of the frozen-core CCSD(T)/CBS limit. The result is consistent with the both CCSD(T)-F12b/CBS(Q-5)+M (with and without scaled triple) results. Finally for calculating potential energy surface, use CCSD(T)-F12b/CBS(Q-5)+M approach which gives here a value -205.835 cm$^{-1}$ with the estimated error 0.4 cm$^{-1}$.

Since the value -205.835 cm$^{-1}$ (CCSD(T)-F12b/CBS(Q-5)+M) is within the error range of my benchmark value ($-206.080 \pm 0.4$), I can assume that it reproduces the benchmark value and the expected error should be below 0.4 cm$^{-1}$. It is very unlikely that the true value for CBS limit of CCSD(T) method is somehow above -206.080 cm$^{-1}$ (since the trend for CBS(T-Q)+M, CBS(Q-5)+M, CBS(5-6)+M goes from smaller values to bigger ones).

There is one more possibility to estimate the interaction energy due to fact that the scaled-triples result for the basis set with midbond functions tend to overshoot the interaction energy[11]. The CCSD(T)-F12b (without scaling) results tend to somewhat underestimate it. Therefore good estimation can be provided by averaged value of CCSD(T)-F12b methods, with and without scaled triples, which gives value -206.118 cm$^{-1}$. This value is again within the error range of my best estimations for CCSD(T)/CBS(5-6)+M and CCSD(T)-F12b/CBS(Q-5)+M.

**Table 4-1** Interaction energies (in cm$^{-1}$) calculated by CCSD(T) and CCSD(T)-F12b (both with and without scaled triple excitations) methods for diffent basis sets.

| Basis | CCSD(T)-F12b | | CCSD(T)-F12b (scaled triples) | | CCSD(T) | |
|---|---|---|---|---|---|---|
| | E(X)[*] | E(X)-E(X-1) | E(X) | E(X)-E(X-1) | E(X) | E(X)-E(X-1) |
| AVDZ+M | -185.845 | | -197.026 | | -158.961 | |
| AVTZ+M | -201.518 | -15.67 | -205.518 | -8.68 | -193.692 | -34.73 |
| AVQZ+M | -204.624 | -3.11 | -206.231 | -0.53 | -202.361 | -8.67 |
| AV5Z+M | -205.062 | -0.44 | -206.161 | 0.07 | -203.921 | -1.56 |
| AV6Z+M | | | | | -204.683 | -0.76 |
| *CBS(T-Q)+M* | *-207.061* | | *-206.788* | | *-209.320* | |
| *CBS(Q-5)+M* | *-205.835* | *1.23* | *-206.401* | *0.53* | *-206.487* | *2.83* |
| *CBS(5-6)+M* | | | | | *-206.080* | *0.41* |
| AVDZ | -171.656 | | -181.128 | | -122.888 | |
| AVTZ | -198.449 | -26.79 | -202.582 | -21.45 | -183.349 | -60.46 |
| AVQZ | -203.723 | -5.27 | -205.372 | -2.79 | -199.442 | -16.09 |
| AV5Z | -204.680 | -0.96 | -205.821 | -0.45 | -202.779 | -3.34 |
| AV6Z | | | | | -204.096 | -1.32 |
| *CBS(T-Q)* | *-207.746* | | *-207.583* | | *-210.647* | |
| *CBS(Q-5)* | *-205.984* | *1.76* | *-206.593* | *0.99* | *-207.019* | *3.63* |
| *CBS(5-6)* | | | | | *-206.240* | *0.779* |

---

[*] E(X) stands for the interaction energy calculated in the AV*X*Z basis set.

## 4.1.2 CCSDT and CCSDT(Q) contributions

Since nowadays inclusion of higher excitations beyond CCSD(T) method are available, I tested the inclusion of full triple excitations CCSDT, of noniterative quadruple excitations CCSDT(Q), and of full quadruple excitations CCSDTQ.

Let's consider four quantities related to various levels of excitations in the coupled-cluster method:

$$\delta E_{int}^{T-(T)} = E_{int}^{CCSDT} - E_{int}^{CCSD(T)},$$

$$\delta E_{int}^{(Q)-T} = E_{int}^{CCSDT(Q)} - E_{int}^{CCSDT},$$

$$\delta E_{int}^{(Q)-(T)} = E_{int}^{CCSDT(Q)} - E_{int}^{CCSD(T)}, \text{ and}$$

$$\delta E_{int}^{Q-(Q)} = E_{int}^{CCSDTQ} - E_{int}^{CCSDT(Q)}$$

The values of these quantities are presented in Table 4-2. $\delta E_{int}^{Q-(Q)}$ can be ignored entirely and assumed that only 0.01 cm$^{-1}$ for the uncertainty comes from the neglect of full quadruples and all higher coupled-cluster excitations.

The contribution of the midbond functions is very small, as well, especially if we consider the computational times increases strongly with number of functions.

The sum $\delta E_{int}^{(Q)-(T)}$ converges faster with the basis set than the $\delta E_{int}^{T-(T)}$ and $\delta E_{int}^{(Q)-T}$ terms separately, but one should exercise a great deal of caution. The estimated CBS limit for all three quantities is as follows: $\delta E_{int}^{T-(T)} = -1.43 \pm 0.25$ cm$^{-1}$ (the CBS(Q-5) result and the estimated error is the difference between CBS(Q-5) and AV5Z), $\delta E_{int}^{(Q)-T} = -2.15 \pm 0.16$ cm$^{-1}$ (the CBS(T-Q) result and the estimated error is the difference between CBS(T-Q) and AVQZ) and a sum of the above two values

$\delta E_{int}^{(Q)-(T)}$ =-3.58 ± 0.30 (the estimated error is calculated by $\sqrt{0.25^2 + 0.16^2}$). We can

compare it to another CBS limit $\delta E_{int}^{(Q)-(T)}$/CBS(T-Q) value of -3.72 ± 0.20cm$^{-1}$ (the

estimated error is the difference between CBS(T-Q) and AVQZ). The two estimates of

$\delta E_{int}^{(Q)-(T)}$ are perfectly consistent.

Here I tried to find a better way of calculating the $\delta E_{int}^{(Q)-(T)}$ correction without

a need to compute CCSDT(Q) for any basis larger than AVDZ+M/AVDZ because

computational cost is too big for computing it for the whole potential energy surface.

The difference $\delta E_{int}^{T-(T)}$ decreases as the basis set size increases while the

$\delta E_{int}^{(Q)-T}$ difference shows opposite trend. The best approach in this case would be to

calculate $\delta E_{int}^{T-(T)}$ in a larger AVTZ basis set and $\delta E_{int}^{(Q)-T}$ in AVDZ. This gives a

value of -3.55 cm$^{-1}$ which is a very reasonable estimate. The error which comes from

different basis sets might cancel each other. I can also try to calculate $\delta E_{int}^{T-(T)}$

/AVTZ+M and $\delta E_{int}^{(Q)-T}$/AVDZ+M which gives -3.46 cm$^{-1}$ but this does not seem like

any improvement. Also adding midbond functions to a basis set as small as AVDZ

might result in an unbalanced basis that actually performs worse. In the view of the

benchmark results with the uncertainty of about 0.3 cm$^{-1}$ gives the estimate of the

post-CCSD(T) effects to -3.55 ± 0.30 cm$^{-1}$.

**Table 4-2** Comparison of the interaction energy for the CCSD(T), CCSDT, CCSDT(Q), and CCSDTQ calculations and values of the $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, $\delta E_{int}^{(Q)-(T)}$, and $\delta E_{int}^{Q-(Q)}$ contributions. All energies are given in cm$^{-1}$.

| Basis | CCSD(T) | | CCSDT | | | CCSDT(Q) | | | | CCSDTQ | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | E(X)$^*$ | E(X)-E(X-1) | E(X) | E(X)-E(X-1) | $\delta E_{int}^{T-(T)}$ | E(X) | E(X)-E(X-1) | $\delta E_{int}^{(Q)-T}$ | $\delta E_{int}^{(Q)-(T)}$ | E(X) | $\delta E_{int}^{Q-(Q)}$ |
| AVDZ+M | -158.961 | | -161.971 | | **-3.01** | -163.162 | | **-1.19** | **-4.20** | -163.153 | **0.01** |
| AVTZ+M | -193.692 | -34.73 | -195.966 | -34.00 | **−2.27** | -197.792 | -34.63 | **-1.83** | **-4.10** | | |
| AVQZ+M | -202.361 | -8.67 | -204.169 | -8.20 | **-1.81** | -206.175 | -8.38 | **-2.01** | **-3.81** | | |
| AV5Z+M | -203.921 | -1.56 | | | | | | | | | |
| AV6Z+M | -204.683 | -0.76 | | | | | | | **-4.06** | | |
| *CBS(T-Q)M* | *-209.320* | | | | ***-1.47*** | | | ***-2.14*** | ***-3.61*** | | |
| *CBS(Q-5)M* | *-206.487* | *2.83* | | | | | | | | | |
| *CBS(5-6)M* | *-206.080* | *0.41* | | | | | | | | | |

---

$^*$ E(X) stands for the interaction energy calculated in the AV$X$Z+M basis set.

**Table 4-2** (continued)

| | CCSD(T) | | CCSDT | | | CCSDT(Q) | | | | CCSDTQ | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Basis | $E(X)^{*}$ | $E(X)-E(X-1)$ | $E(X)$ | $E(X)-E(X-1)$ | $\delta E_{int}^{T-(T)}$ | $E(X)$ | $E(X)-E(X-1)$ | $\delta E_{int}^{(Q)-T}$ | $\delta E_{int}^{(Q)-(T)}$ | $E(X)$ | $\delta E_{int}^{Q-(Q)}$ |
| AVDZ | -122.888 | | -125.611 | | **-2.72** | -126.733 | | **-1.12** | **-3.84** | -126.743 | **-0.01** |
| AVTZ | -183.349 | -60.46 | -185.775 | -60.16 | **-2.43** | -187.549 | -60.82 | **-1.77** | **-4.20** | | |
| AVQZ | -199.442 | -16.09 | -201.369 | -15.59 | **-1.93** | -203.361 | -15.81 | **-1.99** | **-3.92** | | |
| AV5Z | -202.779 | -3.34 | -204.462 | -3.09 | **-1.68** | | | | | | |
| AV6Z | -204.096 | -1.32 | | | | | | | | | |
| *CBS(T-Q)* | *-210.647* | | *-212.392* | | ***-1.75*** | | | ***-2.15*** | ***-3.72*** | | |
| *CBS(Q-5)* | *-207.019* | *3.63* | *-208.446* | *3.98* | ***-1.43*** | | | | | | |

---

[*] E(X) stands for the interaction energy calculated in the AV*X*Z basis set.

## 4.1.3 CCSD(T) core electron correction

Here I denote the difference between interaction energy calculated by the all electron-correlated CCSD(T) method and frozen-core electron CCSD(T) method as $\delta E_{int}^{AE}$. Values of the $\delta E_{int}^{AE}$, all electron-correlated CCSD(T), and frozen-core electron CCSD(T) are presented in Table 4-3. Based on the results presented in the table I can draw 3 conclusions:

First, the addition of midbond functions does not make a significant enhancement. Second, standard AV$X$Z bases converge very slow.

Third, the agreement between ACV$X$Z and AWCV$X$Z results is very good. The latter basis set can be viewed as a little bit better.

The core-valence electron correlation should be more important for interaction energy calculations than the core-core electron correlation, and AWCV$X$Z emphasizes better description of the core-valence electron correlation. Therefore CCSD(T) with the (AWCVTZ, AWCVQZ) extrapolation gives an excellent estimate $-1.80 \pm 0.20$ cm$^{-1}$ for the core electron correlation. The esitmate error comes from the observation that the difference between the CBS(AWCVTZ, AWCVQZ) and CBS(AWCVQZ, AWCV5Z) is 0.08 cm$^{-1}$, therefor the best available value is $-1.88 \pm 0.08$ cm$^{-1}$. So using value $-1.80$ the biggest possible error should be smaller than 0.20 cm$^{-1}$. It is worth to note that for AWCV5Z basis set the correction is 1.74 cm$^{-1}$. All these values are again in the range of estimated error ($1.80 \pm 0.20$ cm$^{-1}$). No result in the regular AV$X$Z bases is even close to this accuracy.

**Table 4-3** The electron-correlated correction $\delta E_{int}^{AE}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T) | CCSD(T)-AE (All electron-correlated) | $\delta E_{int}^{AE}$ |
|---|---|---|---|
| ACVTZ+M | -192.010 | -193.304 | **-1.29** |
| ACVQZ+M | -201.212 | -202.750 | **-1.54** |
| *CBS(T-Q)+M* | *-208.210* | *-209.924* | ***-1.72*** |
| AWCVTZ+M | -191.605 | -192.972 | **-1.37** |
| AWCVQZ+M | -201.297 | -202.895 | **-1.60** |
| *CBS(T-Q)+M* | *-208.339* | *-210.139* | ***-1.80*** |
| ACVTZ | -181.773 | -183.057 | **-1.29** |
| ACVQZ | -198.284 | -199.832 | **-1.55** |
| ACV5Z | -202.480 | -204.188 | **-1.71** |
| *CBS(T-Q)* | *-209.762* | *-211.502* | ***-1.74*** |
| *CBS(Q-5)* | *-206.662* | *-208.537* | ***-1.88*** |
| AWCVTZ | -181.481 | -182.804 | **-1.32** |
| AWCVQZ | -198.382 | -199.979 | **-1.60** |
| AWCV5Z | -202.539 | -204.276 | **-1.74** |
| *CBS(T-Q)* | *-209.872* | *-211.669* | ***-1.80*** |
| *CBS(Q-5)* | *-206.637* | *-208.521* | ***-1.88*** |
| AVTZ | -183.349 | -183.238 | **0.11** |
| AVQZ | -199.442 | -199.412 | **0.03** |
| AV5Z | -202.779 | -203.609 | **-0.83** |
| AV6Z | -204.096 | -205.241 | **-1.15** |
| *CBS(T-Q)* | *-209.320* | *-210.987* | ***-1.67*** |
| *CBS(Q-5)* | *-206.487* | *-208.918* | ***-2.43*** |
| *CBS(5-6)* | *-206.080* | *-207.705* | ***-1.63*** |

## 4.1.4 Relativistic correction

Here I denote the difference between interaction energy calculated by the all electron-correlated, relativistic corrected CCSD(T) method and all electron-correlated CCSD(T) method as $\delta E_{int}^{rel}$. The relativistic corrections are presented in Table 4-4. The additional "d" in front of each abbreviated basis set notation means the basis set is decontracted. The relativistic correction is more important for inner-shell electrons, so midbond functions were not included. Most reasonable basis sets give very close results, but this is not true in the case of AV$X$Z sets. Sufficient flexibility of the basis set in the large-exponent range is needed. Such flexibility is not provided by the AV$X$Z basis set and that explains here the slow convergence with the size of basis set. The ACV$X$Z and AWCV$X$Z bases are better because they provide additional functions with large exponents, but decontracted basis sets have even more flexibility and should be still better.

Decontracted AV$X$Z bases are easier to work with than decontracted ACV$X$Z bases where the convergence problem makes calculation difficult. This problem rises from the additional tight exponents in dACV$X$Z bases which create nearly linear dependencies. For these reasons, use the dAVQZ basis set which gives the relativistic correction of $1.06 \pm 0.05$ cm$^{-1}$. The uncertainty is assumed somewhat arbitrarily, the data presented in the Table 4-4 suggest that the error should be much smaller.

**Table 4-4** The relativistic correction ( $\delta E_{int}^{rel}$ ). The energies are given in cm$^{-1}$ .

| Basis | CCSD(T)-AE (With all electron correlated) | CCSD(T) (With relativistic correction, all electrons are correlated ) | $\delta E_{int}^{rel}$ |
|---|---|---|---|
| ACVTZ | -183.057 | -182.121 | **0.94** |
| ACVQZ | -199.832 | -198.916 | **0.92** |
| ACV5Z | -204.188 | -203.242 | **0.95** |
| *CBS(T-Q)* | *-211.502* | *-210.592* | *0.91* |
| *CBS(Q-5)* | *-208.537* | *-207.615* | *0.92* |
| dACVTZ | -185.016 | -183.941 | **1.07** |
| dACV5Z[*] | -204.185 | -203.131 | **1.05** |
| AWCVTZ | -182.804 | -181.841 | **0.96** |
| AWCVQZ | -199.979 | -199.058 | **0.92** |
| AWCV5Z | -204.276 | -203.330 | **0.95** |
| *CBS(T-Q)* | *-211.669* | *-210.746* | *0.92* |
| *CBS(Q-5)* | *-208.521* | *-207.596* | *0.92* |

---

[*] dACVQZ basis set result did not converge.

**Table 4-4** (continued)

| Basis | CCSD(T)-AE (With all electron correlated) | CCSD(T) (With relativistic correction, all electrons are correlated ) | $\delta E_{int}^{rel}$ |
|---|---|---|---|
| AVTZ | -183.238 | -182.863 | **0.38** |
| AVQZ | -199.412 | -199.185 | **0.23** |
| AV5Z | -203.609 | -203.119 | **0.49** |
| AV6Z | -205.241 | -204.299 | **0.94** |
| *CBS(T-Q)* | *-210.987* | *-210.674* | *0.31* |
| *CBS(Q-5)* | *-208.918* | *-208.607* | *0.31* |
| *CBS(5-6)* | *-207.705* | *-206.974* | *0.73* |
| dAVTZ | -185.311 | -184.233 | **1.08** |
| dAVQZ | -199.676 | -198.615 | **1.06** |
| dAV5Z | -203.488 | -202.433 | **1.06** |
| *CBS(T-Q)* | *-211.152* | *-210.091* | *1.06* |
| *CBS(Q-5)* | *-208.088* | *-207.028* | *1.06* |

## 4.1.5 Total energy

Adding all the contributions to the interaction energy and calculating error by taking the square root of the sum of the squares of the respective uncertainty, we estimated interaction energy for $-210.13 \pm 0.54$ cm$^{-1}$ at the $\{R, \theta_1, \theta_2, \phi, r, s\} = \{6.5, 90, 0, 0, 2.409, 1.449\}$ geometry. A linear addition of uncertainties would give $\pm 0.95$ cm$^{-1}$, but that would be too conservative as different errors act in fairly random way. In the next section, I tried to verify that the theory level and basis set mentioned previously can be applied to calculating the whole potential energy surface by repeating similar calculations of interaction energies at several different geometries.

## 4.2 Interaction energy at general geometries

The following discussed geometries are all with fixed $H_2$ and HCl bond distance at $s=1.449$ bohr, $r=2.409$ bohr, respectively.

## 4.2.1 $\{R, \theta_1, \theta_2, \phi\} = \{6.5, 90, 180, 0\}$

CCSD(T) and CCSD(T)-F12b interaction energies and convergence behavior of different basis sets are shown in Table 4-5. CCSD(T)-F12b results show much better convergence than CCSD(T)'s. CCSD(T)/CBS(5-6)+M and CCSD(T)-F12b/CBS(Q-5)+M values are perfectly consistent. The estimated error can be chosen from the difference between CCSD(T)-F12b/CBS(Q-5)+M and CCSD(T)-F12b/AV5Z+M, which gives 0.173 cm$^{-1}$. The estimated error at this geometries is smaller than the error at $\{R, \theta_1, \theta_2, \phi\} = \{6.5, 90, 0, 0\}$ which is 0.4 cm$^{-1}$. Here CCSD(T)-F12b/CBS(Q-5)+M gives very accurate result.

The core electron correction $\delta E_{int}^{AE}$ varies with AWCV$X$Z basis set size but it is

all together small. Results are presented in Table 4-6. All basis gives close results. The estimated errors will be less than 0.1 cm$^{-1}$. Here I choose the CBS(T-Q) value 0.03 cm$^{-1}$ if taking computational time into account.

$\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{(Q)-(T)}$ values are presented in Table 4-7. The best estimate can be chosen from the summation of $\delta E_{int}^{(Q)-T}$/CBS(T-Q), -0.97, and $\delta E_{int}^{T-(T)}$/CBS(Q-5), -0.61, which gives -1.58. Taking computational time into account, the largest basis calculation that can be done are $\delta E_{int}^{T-(T)}$/AVTZ and $\delta E_{int}^{(Q)-T}$/AVDZ. It gives -0.93-0.24=-1.35. The estimated error would be below 0.35 cm$^{-1}$. This estimate also covers the $\delta E_{int}^{(Q)-(T)}$ / CBS(T-Q) result (-1.69 cm$^{-1}$).

The relativistic correction $\delta E_{int}^{rel}$ are presented in Table 4-8. Like core correction results, all dAV$X$Z sets gives close results. Here I choose the dAVQZ result -1.2. The estimated errors will be less than 0.1 cm$^{-1}$.

From the above discussion, I calculate the total interaction energy as a sumation of CCSD(T)-F12b/CBS(Q-5)+M, $\delta E_{int}^{AE}$/CBS(T-Q), $\delta E_{int}^{T-(T)}$/AVTZ, $\delta E_{int}^{(Q)-T}$/AVDZ, and $\delta E_{int}^{rel}$/dAVQZ. It gives -99.86 cm$^{-1}$. The uncertainty is calculated by the square root of the sum of the squares of the respective errors, which gives 0.39 cm$^{-1}$.

**Table 4-5** CCSD(T) and CCSD(T)-F12b interaction energies (in cm$^{-1}$) and convergence behavior of different basis sets.

| Basis | CCSD(T)-F12b | | CCSD(T) | |
|---|---|---|---|---|
| | E(X)[*] | E(X)-E(X-1) | E(X) | E(X)-E(X-1) |
| AVDZ+M | -88.307 | | -80.935 | |
| AVTZ+M | -96.767 | -8.46 | -94.428 | -13.49 |
| AVQZ+M | -97.526 | -0.76 | -96.381 | -1.95 |
| AV5Z+M | -97.517 | 0.01 | -97.143 | -0.76 |
| AV6Z+M | | | -97.305 | -0.16 |
| *CBS(T-Q)+M* | *-97.963* | | *-97.181* | |
| *CBS(Q-5)+M* | *-97.344* | *0.62* | *-97.154* | *0.03* |
| *CBS(5-6)+M* | | | *-97.347* | *-0.19* |

**Table 4-6** The core electron correction $\delta E_{int}^{AE}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T) | CCSD(T) (All electron-correlated) | $\delta E_{int}^{AE}$ |
|---|---|---|---|
| AWCVTZ | -85.351 | -85.507 | -0.16 |
| AWCVQZ | -93.210 | -93.258 | -0.05 |
| AWCV5Z | -95.408 | -95.395 | 0.01 |
| *CBS(T-Q)* | *-98.710* | *-98.680* | *0.03* |
| *CBS(Q-5)* | *-97.771* | *-97.694* | *0.08* |

---

[*] E(X) stands for the interaction energy calculated in the AV*X*Z+M basis set.

**Table 4-7** The values of $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{(Q)-(T)}$ contributions. The energies are given in cm$^{-1}$.

| | CCSD(T) | | CCSDT | | | CCSDT(Q) | | | |
|---|---|---|---|---|---|---|---|---|---|
| Basis | E(X)[*] | E(X)-E(X-1) | E(X) | E(X)-E(X-1) | $\delta E_{int}^{T-(T)}$ | E(X) | E(X)-E(X-1) | $\delta E_{int}^{(Q)-T}$ | $\delta E_{int}^{(Q)-(T)}$ |
| AVDZ | -55.565 | | -56.773 | | **-1.21** | -57.194 | | **-0.42** | **-1.63** |
| AVTZ | -84.382 | -28.82 | -85.314 | -28.54 | **-0.93** | -86.043 | -28.8495 | **-0.73** | **-1.66** |
| AVQZ | -92.851 | -8.47 | -93.666 | -8.35 | **-0.81** | -94.532 | -8.4886 | **-0.87** | **-1.68** |
| AV5Z | -95.307 | -2.46 | -96.023 | -2.36 | **-0.72** | | | | |
| *CBS(T-Q)* | | | | | ***-0.73*** | | | ***-0.97*** | ***-1.69*** |
| *CBS(Q-5)* | | | | | ***-0.61*** | | | | |

---

[*] E(X) stands for the interaction energy calculated in the AV$X$Z basis set.

**Table 4-8** The relativistic correction $\delta E_{int}^{rel}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T)-AE (With all electron correlated) | CCSD(T) (With relativistic correction, all electrons are crrelated ) | $\delta E_{int}^{rel}$ |
|---|---|---|---|
| dAVTZ | -84.870 | -86.112 | -1.24 |
| dAVQZ | -93.204 | -94.405 | -1.20 |
| dAV5Z | -95.748 | -96.938 | -1.19 |

## 4.2.2 {R, $\theta_1$, $\theta_2$, $\phi$}={7, 0, 0, 0}

CCSD(T) and CCSD(T)-F12b interaction energies and convergence behavior of different basis sets are shown in Table 4-9. Both CCSD(T)-F12b and CCSD(T) results converge well. CCSD(T)/CBS(5-6)+M and CCSD(T)-F12b/CBS(Q-5)+M values are consistent, and the difference is around 0.12 cm$^{-1}$. The estimated uncertainty would be much less than 0.2 cm$^{-1}$.

The core electron correction $\delta E_{int}^{AE}$ varied with the AWCV$X$Z basis sets size. Results are presented in Table 4-10. Again, all basis gives close results. The estimated errors would be less than 0.1 cm$^{-1}$.

$\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{(Q)-(T)}$ values are presented in Table 4-11. The best estimates can be chosen from the summation of $\delta E_{int}^{(Q)-T}$/CBS(T-Q), -1.93, and $\delta E_{int}^{T-(T)}$/CBS(Q-5), -1.78, which gives -3.71. Again taking computational time into account, the largest basis calculation that can be done for all grid points are $\delta E_{int}^{T-(T)}$ /AVTZ and $\delta E_{int}^{(Q)-T}$ /AVDZ. It gives -2.66-0.80=-3.46. Comparing these two values provides the error estimating less than 0.35 cm$^{-1}$. This estimate also covers the $\delta E_{int}^{(Q)-(T)}$/ CBS(T-Q) result (-3.80 cm$^{-1}$).

The relativistic correction $\delta E_{int}^{rel}$ is presented in Table 4-12. All dAV$X$Z sets give almost identical results. The estimated errors would be less than 0.1 cm$^{-1}$.

Therefor the total interaction energy is [CCSD(T)-F12b/CBS(Q-5)+M]+[$\delta E_{int}^{AE}$ /CBS(T-Q)]+[ $\delta E_{int}^{T-(T)}$ /AVTZ]+[ $\delta E_{int}^{(Q)-T}$ /AVDZ]+[ $\delta E_{int}^{rel}$ /dAVQZ]. The total interaction energy at {R, $\theta_1$, $\theta_2$, $\phi$}={7, 0, 0, 0} is 138.842 cm$^{-1}$. The uncertainty is

calculated by the square root of the sum of the squares of the respective errors, this gives $0.39 \text{ cm}^{-1}$.

**Table 4-9** CCSD(T) and CCSD(T)-F12b interaction energies (in cm$^{-1}$).

| Basis | CCSD(T)-F12b | | CCSD(T) | |
|---|---|---|---|---|
| | E(X)$^*$ | E(X)-E(X-1) | E(X) | E(X)-E(X-1) |
| AVDZ+M | 152.011 | | 155.479 | |
| AVTZ+M | 144.213 | -7.80 | 144.460 | -11.02 |
| AVQZ+M | 143.280 | -0.93 | 143.366 | -1.09 |
| AV5Z+M | 143.258 | -0.02 | 143.212 | -0.15 |
| AV6Z+M | | | 143.176 | -0.04 |
| *CBS(T-Q)+M* | *142.749* | | *142.818* | |
| *CBS(Q-5)+M* | *143.292* | *0.54* | *143.341* | *0.52* |
| *CBS(5-6)+M* | | | *143.170* | *-0.17* |

**Table 4-10** The core electron correction $\delta E_{int}^{AE}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T) | CCSD(T) (All electron-orrelated) | $\delta E_{int}^{AE}$ |
|---|---|---|---|
| AWCVTZ | 150.394 | 149.926 | -0.47 |
| AWCVQZ | 145.173 | 144.713 | -0.46 |
| AWCV5Z | 143.734 | 143.291 | -0.45 |
| *CBS(T-Q)* | *142.527* | *142.071* | *-0.46* |
| *CBS(Q-5)* | *142.759* | *142.325* | *-0.43* |

---

$^*$ E(X) stands for the interaction energy calculated in the AV$X$Z+M basis set.

**Table 4-11** The values of $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{(Q)-(T)}$ contributions. The energies are given in cm$^{-1}$.

| | CCSD(T) | | CCSDT | | | CCSDT(Q) | | | |
|---|---|---|---|---|---|---|---|---|---|
| Basis | E(X)[*] | E(X)-E(X-1) | E(X) | E(X)-E(X-1) | $\delta E_{int}^{T-(T)}$ | E(X) | E(X)-E(X-1) | $\delta E_{int}^{(Q)-T}$ | $\delta E_{int}^{(Q)-(T)}$ |
| AVDZ | 163.206 | | 159.969 | | **-3.24** | 159.172 | | **-0.80** | **-4.03** |
| AVTZ | 150.181 | -13.03 | 147.517 | -12.45 | **-2.66** | 145.949 | -13.22 | **-1.57** | **-4.23** |
| AVQZ | 144.947 | -5.23 | 142.742 | -4.78 | **-2.21** | 140.966 | -4.98 | **-1.78** | **-3.98** |
| AV5Z | 143.682 | -1.27 | 141.686 | -1.06 | **-2.00** | | | | |
| *CBS(T-Q)* | | | | | ***-1.87*** | | | ***-1.93*** | ***-3.80*** |
| *CBS(Q-5)* | | | | | ***-1.78*** | | | | |

---

[*] E(X) stands for the interaction energy calculated in the AV$X$Z basis set.

**Table 4-12**   The relativistic correction $\delta E_{int}^{rel}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T)-AE (With all electron correlated) | CCSD(T) (With relativistic correction, all electrons are correlated ) | $\delta E_{int}^{rel}$ |
|-------|------|------|------|
| dAVTZ | 148.911 | 148.397 | -0.51 |
| dAVQZ | 144.483 | 143.958 | -0.53 |
| dAV5Z | 143.155 | 142.624 | -0.53 |

## 4.2.3 {R, $\theta_1$, $\theta_2$, $\phi$}={7, 90, 90, 90}

CCSD(T) and CCSD(T)-F12b interaction energies and convergence behavior of different basis sets are shown in Table 4-13. Both CCSD(T)-F12b and CCSD(T) results converge. CCSD(T)/CBS(5-6)+M and CCSD(T)-F12b/CBS(Q-5)+M values are consistent, the difference is around 0.1 cm$^{-1}$. The estimated uncertainty would be 0.25 cm$^{-1}$ approximately.

The core electron correction $\delta E_{int}^{AE}$ varied with the AWCV$X$Z basis set and extrapolated results are presented in Table 4-14. Again, all basis gives close results. The estimated errors will be less than 0.05 cm$^{-1}$.

I cannot perform CCSDT and CCSDT(Q) with large basis set (AVQZ, AV5Z) calculations because of the convergence problem which comes from nonplanar geometries. Here I assume conservative estimate of 0.35 cm$^{-1}$ for the uncertainty.

The relativistic correction $\delta E_{int}^{rel}$ is presented in Table 4-15. All dAV$X$Z sets gives almost identical results. The estimated errors will be less than 0.05 cm$^{-1}$.

Therefore, the total interaction energy is [CCSD(T)-F12b/CBS(Q-5)+M]+ [ $\delta E_{int}^{AE}$ /CBS(T-Q)]+[ $\delta E_{int}^{T-(T)}$ /AVTZ (-0.91 cm$^{-1}$)]+[ $\delta E_{int}^{(Q)-T}$ /AVDZ (-0.24 cm$^{-1}$)]+ [ $\delta E_{int}^{rel}$ /dAVQZ]. The total interaction energy at {R, $\theta_1$, $\theta_2$, $\phi$}={7, 90, 90, 90} is -39.20 cm$^{-1}$. The uncertainty is calculated by the square root of the sum of the squares of the respective errors, this gives 0.40 cm$^{-1}$.

**Table 4-13** CCSD(T) and CCSD(T)-F12b interaction energies (in cm$^{-1}$).

| Basis | CCSD(T)-F12b | | CCSD(T) | |
|---|---|---|---|---|
| | E(X)$^*$ | E(X)-E(X-1) | E(X) | E(X)-E(X-1) |
| AVDZ+M | -27.639 | | -20.660 | |
| AVTZ+M | -35.644 | -8.01 | -35.337 | -14.68 |
| AVQZ+M | -36.735 | -1.09 | -36.812 | -1.48 |
| AV5Z+M | -37.179 | -0.44 | -37.170 | -0.36 |
| AV6Z+M | | | -37.316 | -0.15 |
| *CBS(T-Q)M* | *-37.523* | | *-38.114* | |
| *CBS(Q-5)M* | *-37.637* | *-0.11* | *-37.654* | *0.46* |
| *CBS(5-6)M* | | | *-37.540* | *0.11* |

**Table 4-14** The core electron correction $\delta E_{int}^{AE}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T) | CCSD(T) (All electron-orrelated) | $\delta E_{int}^{AE}$ |
|---|---|---|---|
| AWCVTZ | -27.097 | -27.439 | -0.34 |
| AWCVQZ | -33.176 | -33.509 | -0.33 |
| AWCV5Z | -35.530 | -35.852 | -0.32 |
| *CBS(T-Q)* | *-37.889* | *-38.213* | *-0.32* |
| *CBS(Q-5)* | *-38.108* | *-38.420* | *-0.31* |

---

$^*$ E(X) stands for the interaction energy calculated in the AV*X*Z+M basis set.

**Table 4-15** The relativistic correction $\delta E_{int}^{rel}$. The energies are given in cm$^{-1}$.

| Basis | CCSD(T)-AE (With all electron correlated) | CCSD(T) (With relativistic correction, all electrons are crrelated ) | $\delta E_{int}^{rel}$ |
|-------|-------------------------------------------|---------------------------------------------------------------------|------------------------|
| dAVTZ | -27.074 | -27.167 | -0.09 |
| dAVQZ | -33.417 | -33.504 | -0.09 |
| dAV5Z | -35.812 | -35.896 | -0.08 |

# 4.2.4 Summary of the calculations of interaction energies

Results presented in section 4.1 and 4.2 show that the theory level and basis sets mentioned in section 4.1 can be applied to calculating the interaction energy surface at g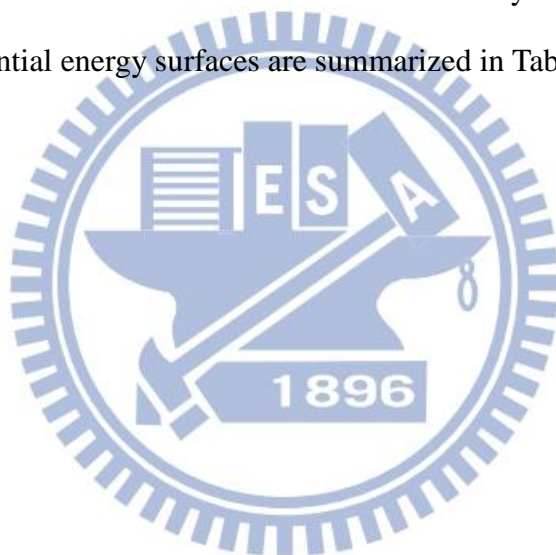rid points of the $H_2$–HCl complex with the total uncertainty of 0.6 cm$^{-1}$. The estimated uncertainty at the different levels of theory and the total uncertainty for 4 chosen geometries are summarized in Table 4-16. The theory level and basis sets used to calculate the potential energy surfaces are summarized in Table 4-17.

**Table 4-16** Estimated uncertainty at the different levels of theory and the total uncertainty for 4 chosen geometries. The energies are given in cm$^{-1}$. Total uncertainties are calculated by the square root of the sum of the squares of the respective uncertainty.

| $\{R,\theta_1,\theta_2,\phi\}$ | $E_{int}^{CCSD(T)-F12b}$ | $\delta E_{int}^{AE}$ | $\delta E_{int}^{T-(T)}+\delta E_{int}^{(Q)-T}$ | $\delta E_{int}^{rel}$ | $E_{int}$ |
|---|---|---|---|---|---|
| $\{6.5,90,0,0\}$ | $-205.835\pm0.4$ | $-1.8\pm0.2$ | $-3.55\pm0.3$ | $1.06\pm0.05$ | $-210.13\pm0.54$ |
| $\{6.5,90,180,0\}$ | $-97.344\pm0.17$ | $0.03\pm0.1$ | $-1.35\pm0.35$ | $-1.2\pm0.1$ | $-99.86\pm0.41$ |
| $\{7,0,0,0\}$ | $143.292\pm0.2$ | $-0.46\pm0.1$ | $-3.46\pm0.35$ | $-0.53\pm0.1$ | $138.84\pm0.43$ |
| $\{7,90,90,90\}$ | $-37.637\pm0.25$ | $-0.32\pm0.05$ | $-1.15\pm0.35$ | $-0.09\pm0.05$ | $-39.20\pm0.44$ |

**Table 4-17** The description of the theory level and basis sets used to calculate potential energy surfaces.

$$E_{int} = E_{int}^{CCSD(T)-F12b} + \delta E_{int}^{AE} + \delta E_{int}^{T-(T)} + \delta E_{int}^{(Q)-T} + \delta E_{int}^{rel}$$

| | Basis set | Theory level |
|---|---|---|
| $E_{int}^{CCSD(T)-F12b}$ | aug-cc-pVQZ+aug-cc-pv5Z+midbond functions+extrapolation | Interaction energy were calculated by CCSD(T)-F12b, frozen core |
| $\delta E_{int}^{AE}$ | aug-cc-pwCVTZ+ aug-cc-pwCVQZ+extrapolation | Correction including all electron correlation, $E_{int}^{CCSD(T)}$(all electron-correlated )$-E_{int}^{CCSD(T)}$(frozen core electron) |
| $\delta E_{int}^{T-(T)}$ | aug-cc-pVTZ | Correction including triple excitations in coupled cluster methods, $E_{int}^{CCSDT} - E_{int}^{CCSD(T)}$, frozen core |
| $\delta E_{int}^{(Q)-T}$ | aug-cc-pVDZ | Correction including quadruple excitations (perturbative) in coupled cluster methods, $E_{int}^{CCSDT(Q)} - E_{int}^{CCSDT}$, frozen core |
| $\delta E_{int}^{rel}$ | decontracted aug-cc-pVQZ | Correction for relativistic effect, $E_{int}^{CCSD(T)}$(relativistic corrected)$-E_{int}^{CCSD(T)}$, all electron-correlated |

## 4.3 The levels of theory and basis sets used to calculate derivatives $f_{ij}(\mathbf{X})$

We already investigated the levels of theory and basis sets which should be used to calculate the interaction energies to reach the desired accuracy at the reference points. Now the question is how to calculate the derivatives $f_{ij}(\mathbf{X})$. In the following tests, 429 grid points were chosen out of 1119 gird points. Those grids were chosen by combining the following values of $R$ (in bohr): 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 10, 11, 12, 15, 20 with the set of 33 unique combinations of the following angles: 0°, 45°, 90° and 135° for $\theta_1$; 0°, 45°, 90°, 135° and 180° for $\theta_2$; and 0°, 45° and 90° for $\phi$. The bond length of H$_2$ was fixed at $s_c$-$h_s$=1.449 bohr. The calculated 4 dimensional potential energy surfaces only include HCl vibration. The vibrationally averaging formula in this case becomes

$$\langle V \rangle_v = \langle \chi_v(r)|V(\mathbf{X},r)|\chi_v(r)\rangle \approx f_0(\mathbf{X}) + f_1(\mathbf{X})(\langle r \rangle_v - r_c) + \frac{1}{2}f_2(\mathbf{X})(\langle r^2 \rangle_v - 2r_c \langle r \rangle_v + r_c^2),$$

$$f_0(\mathbf{X},r_c) = V(\mathbf{X},r_c),$$

$$f_1(\mathbf{X},r_c) = \left.\frac{\partial V}{\partial r}\right|_{r=r_c} \approx \frac{V(\mathbf{X},r_c + h_r) - V(\mathbf{X},r_c - h_r)}{2h_r},$$

$$f_2(\mathbf{X},r_c) = \left.\frac{\partial^2 V}{\partial r^2}\right|_{r=r_c} \approx \frac{V(\mathbf{X},r_c + h_r) - 2V(\mathbf{X},r_c) + V(\mathbf{X},r_c - h_r)}{h_r^2}, \ v = 0.$$

For the derivatives $f_k$, $k = 0, 1, 2$ of 429 points in the grid, I calculated each of them by CCSD(T)-F12b method and included the core electron correction $\delta E_{int}^{AE}$, post

CCSD(T) contributions $\delta E_{int}^{T-(T)} + \delta E_{int}^{(Q)-T}$, and the relativistic correction $\delta E_{int}^{rel}$.

Therefore I generated three different vibrationaly averaged interaction energy surfaces to compare them and test the final approximation proposed by Jankowski[3]:

1. Derivatives $f_0$, $f_1$ and $f_2$ were calculated from the interaction energies derived from CCSD(T)-F12b method with inclusion of all $\delta E_{int}^{AE}$, $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$ and $\delta E_{int}^{rel}$ terms. The four-dimensional vibrationaly averaged interaction energy surface was calculated by the formula presented above. This potential is noted as $\left\langle V^{full} \right\rangle = \left\langle V \right\rangle_0$.

2. Derivatives $f_0$, $f_1$ and $f_2$ were computed from the interaction energies, which derived from CCSD(T)-F12b method plus only core electron correction $\delta E_{int}^{AE}$. The four-dimensional vibrationaly averaged interaction energy surface was calculated by the formula presented above, later corrections $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$ and $\delta E_{int}^{rel}$ terms calculated at $(r_c, s_c)$ reference $H_2$ and HCl distances were added. This potential is noted as

$$\left\langle V^{F12b+AE} \right\rangle = \left\langle V \right\rangle_0' + \delta E_{int}^{T-(T)}\left(r_c, s_c\right) + \delta E_{int}^{(Q)-T}\left(r_c, s_c\right) + \delta E_{int}^{rel}\left(r_c, s_c\right).$$

3. Derivatives $f_0$, $f_1$ and $f_2$ were computed from the interaction energies only by CCSD(T)-F12b method. The four-dimensional vibrationaly averaged interaction energy surface was calculated by the formula presented above and finally all corrections $\delta E_{int}^{AE}$, $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$ and $\delta E_{int}^{rel}$ terms calculated at $(r_c, s_c)$ reference $H_2$ and HCl distances were added. This potential is noted as

$$\left\langle V^{F12b} \right\rangle = \left\langle V \right\rangle_0'' + \delta E_{int}^{AE}\left(r_c, s_c\right) + \delta E_{int}^{T-(T)}\left(r_c, s_c\right) + \delta E_{int}^{(Q)-T}\left(r_c, s_c\right) + \delta E_{int}^{rel}\left(r_c, s_c\right).$$

The statistical data for the differences between the vibrationaly averaged potential interaction energy surfaces calculated in 3 different approaches described above, $|\langle V^{full} \rangle - \langle V^{F12b+AE} \rangle|$ and $|\langle V^{full} \rangle - \langle V^{F12b} \rangle|$, are presented in Table 4-18. In both cases, the biggest deviations occurred in the short $R$ distances. Absolute values of these differences are slightly bigger but absolute values of interaction energies are very big here, as well. Additionally some discrepancy for short $R$ distances might be due to the convergence problem we met at the highly repulsive region. For the purpose of simulating the spectrum of HCl trapped in $pH_2$, the error in the repulsive region is acceptable. After removing 3 extreme cases where these differences are the biggest out of 429 values, the errors and standard deviation are reduced drastically. This phenomenon indicated that the large discrepancy happened in very rare cases in the region where quality of the calculated points are somehow lower. Figure 4-1 also shows over 90% of differences are smaller than 0.1 cm$^{-1}$. In the attractive region of potential energy surface ($\langle V^{full} \rangle < 0$), the maximum deviations occurred at {12, 135, 135, 0} and {6, 90, 0, 0}, but they are in the range not much bigger than 0.1 cm$^{-1}$. The uncertainty coming from approximations used in calculating the interaction energy is as big as 0.6 cm$^{-1}$, therefore errors generated by simplification of the calculations for derivatives ($f_1, f_2$) are much smaller.

These small discrepancies indicate that fairly time-consuming calculations at the higher levels of theory and larger basis sets are only needed for ($\mathbf{X}, r_c$) or ($\mathbf{X}, r_c, s_c$) points. One does not need to apply such high levels to calculate derivatives $f_1$ and $f_2$ or $f_{ij}$, $i + j > 0$. Based on this study about one dimensional averaging, I decide to compute the $f_{ij}$ derivatives from interaction energies obtained from $E_{int}^{CCSD(T)-F12b}$ /CBS(AVQZ, AV5Z), since it produces negligible error.

**Table 4-18** The statistical data for $|\langle V^{full}\rangle - \langle V^{F12b+AE}\rangle|$ and $|\langle V^{full}\rangle - \langle V^{F12b}\rangle|$. The energies are given in cm$^{-1}$.

| | $|\langle V^{full}\rangle - \langle V^{F12b+AE}\rangle|$ | $|\langle V^{full}\rangle - \langle V^{F12b}\rangle|$ |
|---|---|---|
| Average | 0.006 | 0.012 |
| Standard deviation | 0.021 | 0.053 |
| Maximum | 0.240 | 0.666 |
| | at {5.5, 0, 0, 0} | at {5.5, 90, 135, 45} |
| | Without 3 extreme values | |
| | at {5.5, 0, 0, 0}, {5.5, 90, 0, 0} , {6, 0, 0, 0} | at {5.5, 0, 0, 0}, {5.5, 90, 135, 45}, {6, 0, 0, 0}, |
| Average | 0.005 | 0.009 |
| Standard deviation | 0.013 | 0.028 |
| Maximum | 0.137 | 0.312 |
| | at {6.5, 0, 0, 0} | at {5.5, 45, 0, 0} |
| | Attractive part ($V^{full}(\mathbf{X}) < 0$) | |
| Average | 0.003 | 0.004 |
| Standard deviation | 0.008 | 0.012 |
| Maximum | 0.090 | 0.104 |
| | at {12, 135, 135, 0} | at{6, 90, 0, 0} |
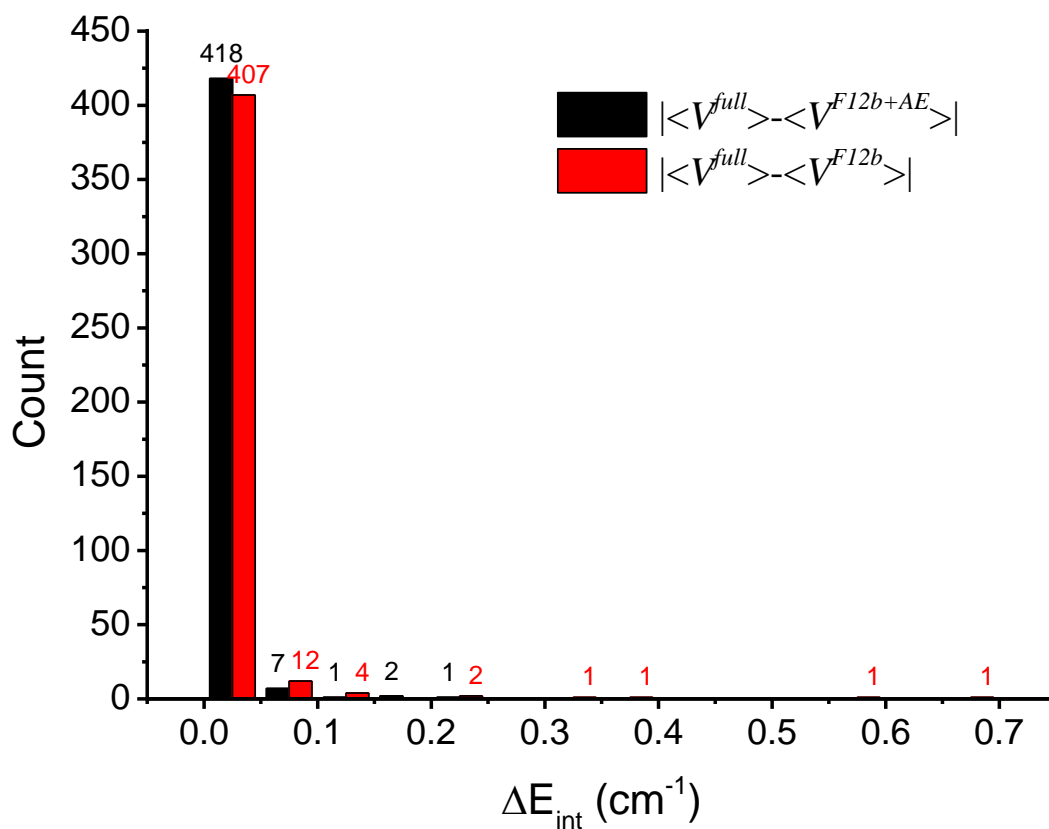
**Figure 4-1** Histogram for $|\langle V^{full}(\mathbf{X})\rangle - \langle V^{F12b+AE}(\mathbf{X})\rangle|$ and $|\langle V^{full}(\mathbf{X})\rangle - \langle V^{F12b}(\mathbf{X})\rangle|$.
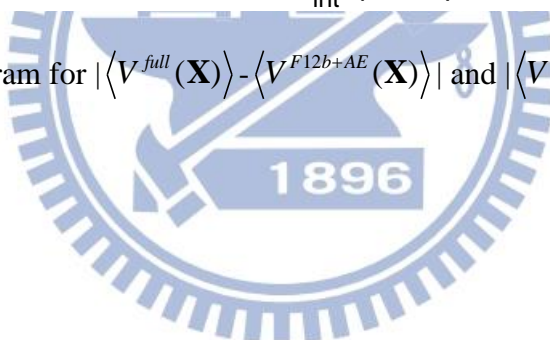
## 4.4 Features of potential energy surfaces

The benefit of the two-step fitting can be illustrated by Figure 4-2. It shows that the shape of the potential energy surfaces along $\theta_2$ is expected to be close to the line shown by CCSD(T)-F12b/AVTZ results. The spline interpolation based on only few points cannot correctly reproduce the shape of whole potential. We produced the shape of the potential $V_{F12b/TZ}$ and be corrected by adding $\Delta V$. The fitting functions for $V_{F12b/TZ}(R, \theta_1, \theta_2, \phi)$ and $\Delta V(R, \theta_1, \theta_2, \phi)$ are linear with respect to the fitting parameters. This makes the fitting procedure relatively simple. The least square approach is used here. In the following sections, the results showing comparison between calculated interaction energies and fitted potential energy surfaces for $V_{F12b/TZ}$, $\Delta V$, and $< V >_{00}$ are presented.
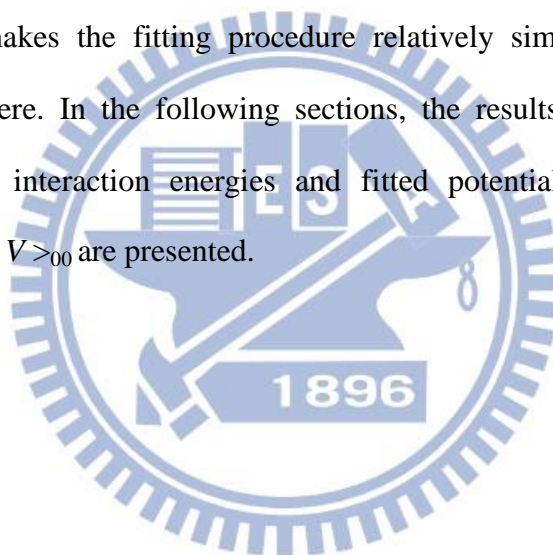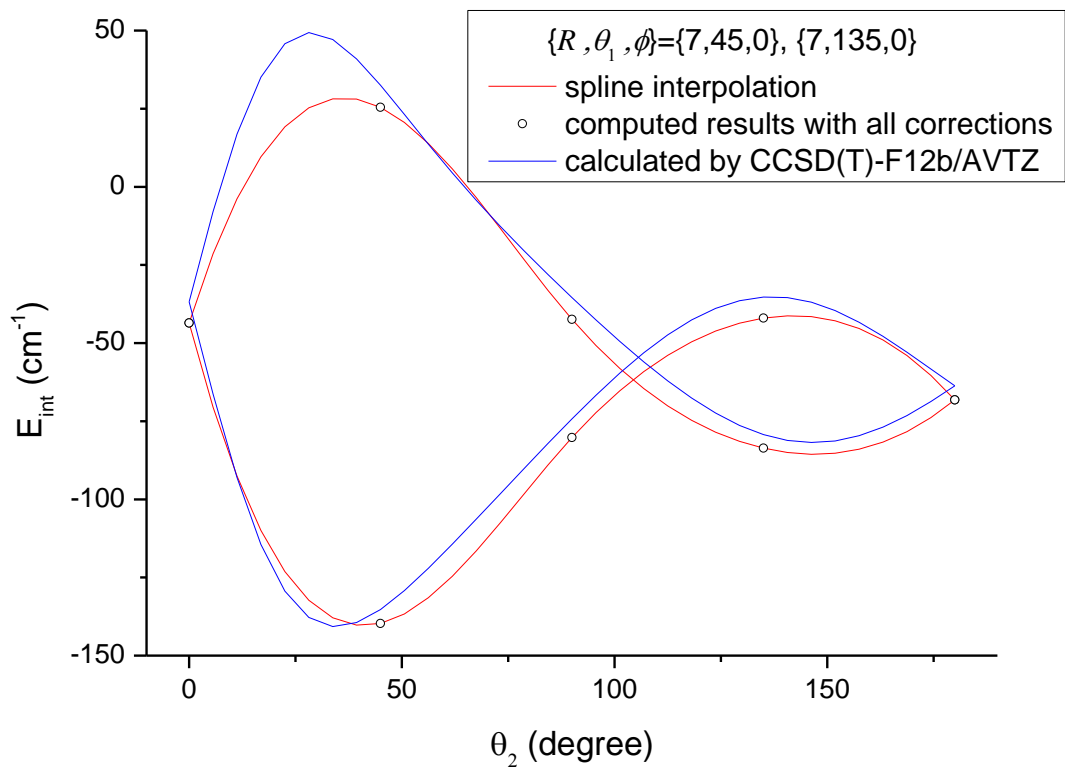
**Figure 4-2**   Shapes of the potential energy surfaces generated by spline interpolation based on calculations including all above mentioned corrections and CCSD(T)-F12b/AVTZ calculation.

## 4.4.1 Fitting results for $V_{F12b/TZ}$

Total of 3003 parameters were used in $V_{F12b/TZ}$ to fit 15459 grid points in $\{R, \ \theta_1, \ \theta_2,$ $\phi\}$ coordinates. Several geometries were chosen to illustrate the results of the fitting in $R, \ \theta_1, \ \theta_2$ and $\phi$ dimensions (Figure 4-3, Figure 4-4, Figure 4-5, Figure 4-6, respectively). The dots in the figures were derived from CCSD(T)-F12b/AVTZ calculations and the lines were corresponding values from the fitting function. The statistical information of the fitting error was shown in Table 4-19.
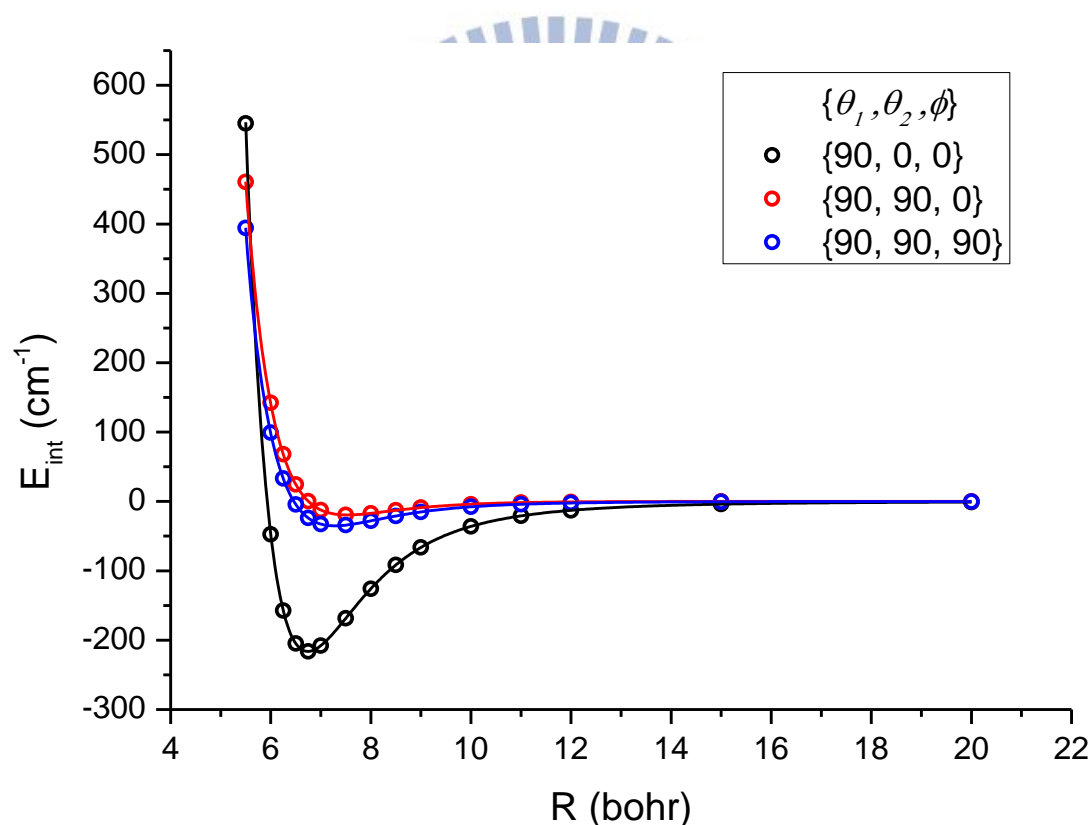


**Figure 4-3** Comparison between calculated interaction energies and fitted potential energy surfaces for $V_{F12b/TZ}$ along R. Dots in the figures were derived from CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.
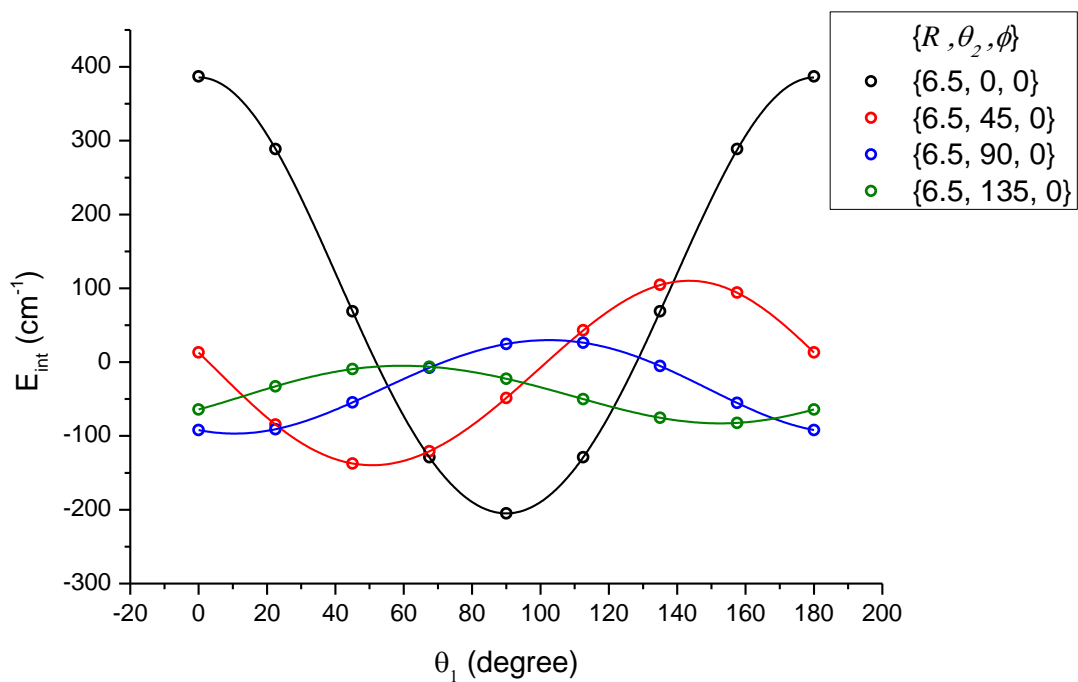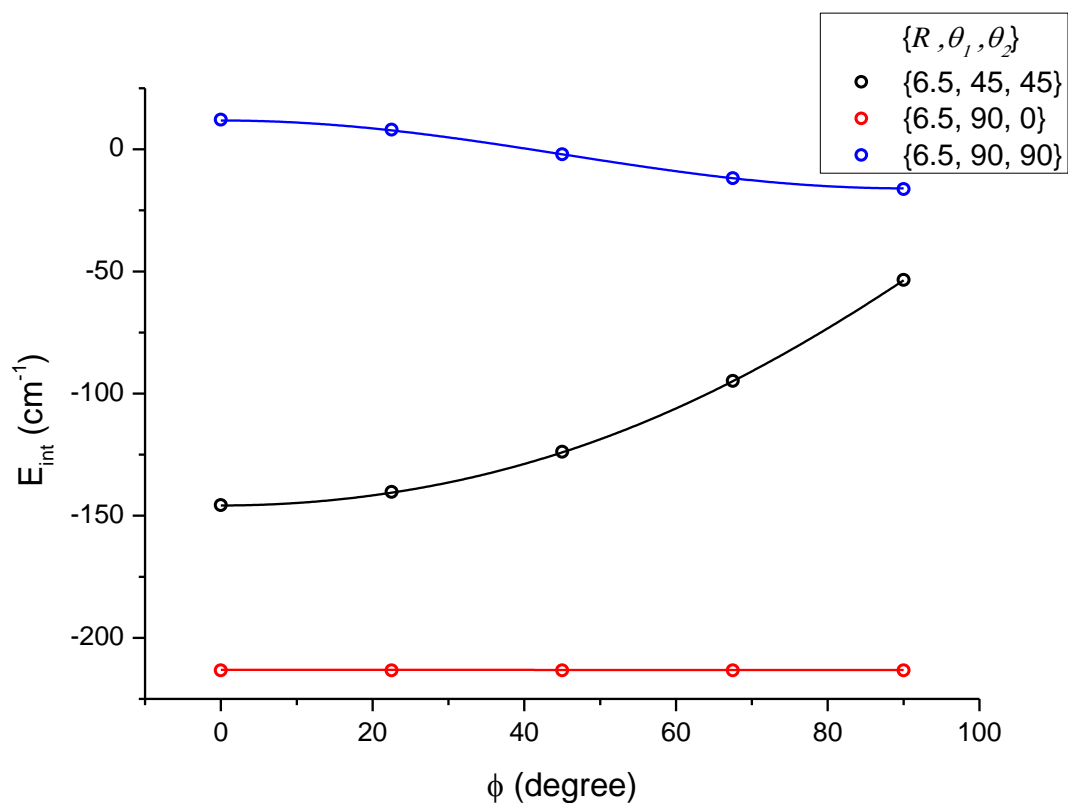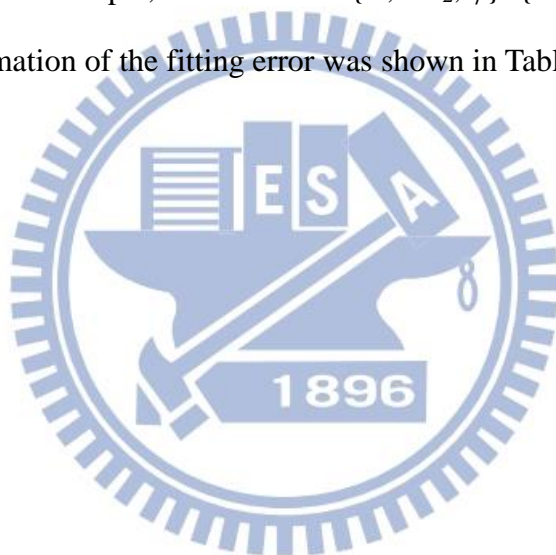
**Figure 4-4** Comparison between calculated interaction energies and fitted potential energy surfaces for $V_{F12b/TZ}$ along $\theta_1$. Dots in the figures were derived from CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.
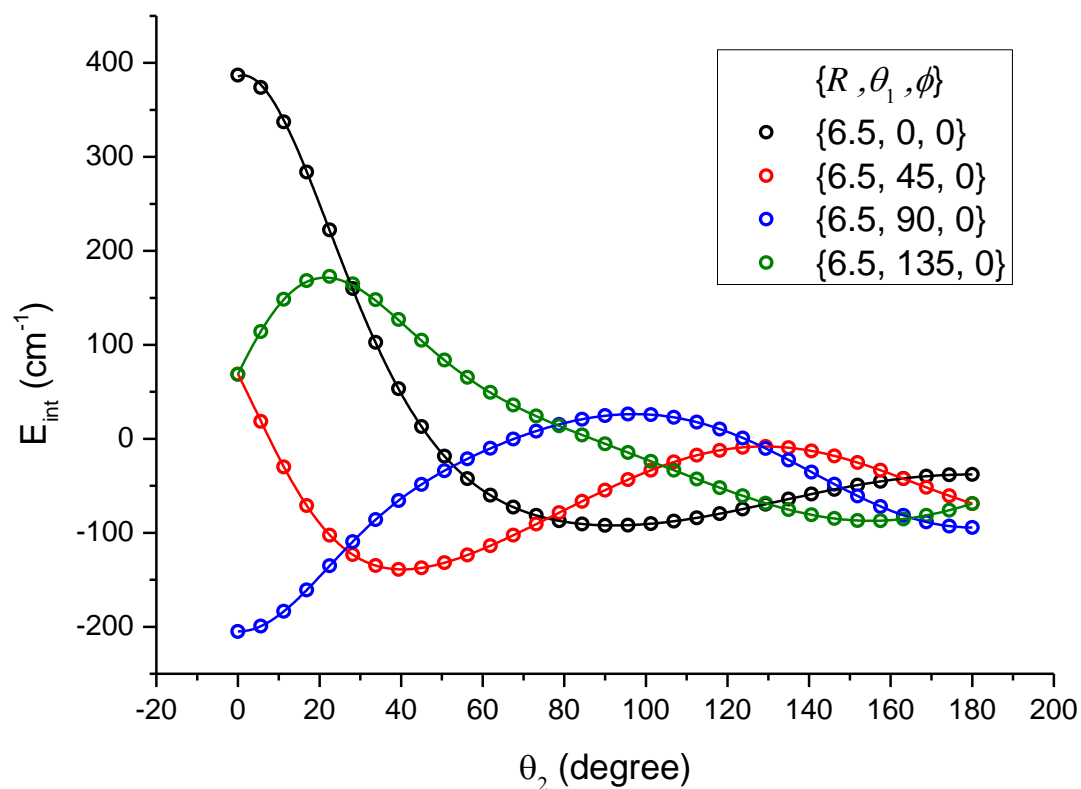
**Figure 4-5** Comparison between calculated interaction energies and fitted potential energy surfaces for $V_{F12b/TZ}$ along $\theta_2$. Dots in the figures were derived from CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.

**Figure 4-6** Comparison between calculated interaction energies and fitted potential energy surfaces for $V_{F12b/TZ}$ along $\phi$. Dots in the figures were derived from CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.

**Table 4-19** The statistical information of the fitting errors for $V_{F12b/TZ}$.

| Errors (cm$^{-1}$) | $V_{F12b/TZ}$ | |
| --- | --- | --- |
| | All points in the grid | Points with negative value of the interaction energy |
| Maximum | 7.89 | <0.110 |
| Mean | 0.082 | 0.004 |
| Standard deviation | 0.024 | 0.0006 |
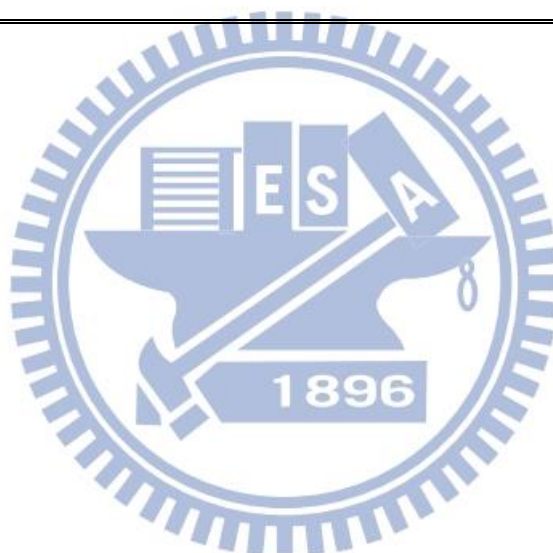
## *4.4.2* **Fitting results for** $\Delta V$

Total of 225 parameters were used in $\Delta V$ to fit 1119 grid points in $\{R,\ \theta_1,\ \theta_2,\ \phi\}$ coordinates. Several geometries were chosen to illustrate the results of the fitting in $R$, $\theta_1$, $\theta_2$ and $\phi$ dimensions (Figure 4-7, Figure 4-8, Figure 4-9, Figure 4-10, respectively). The dots in the figures were derived from the difference between averaged results (see 4.2.4 and 4.3) and CCSD(T)-F12b/AVTZ calculations. Several grid points were discarded in the repulsive region due to the convergence problem in the fitting of $\Delta V$ (For example, four values of $\{R,\ \theta_2,\ \phi\}=\{6.5, 0, 0\}$ in Figure 4-8). The statistical information of the fitting error was shown in Table 4-20.
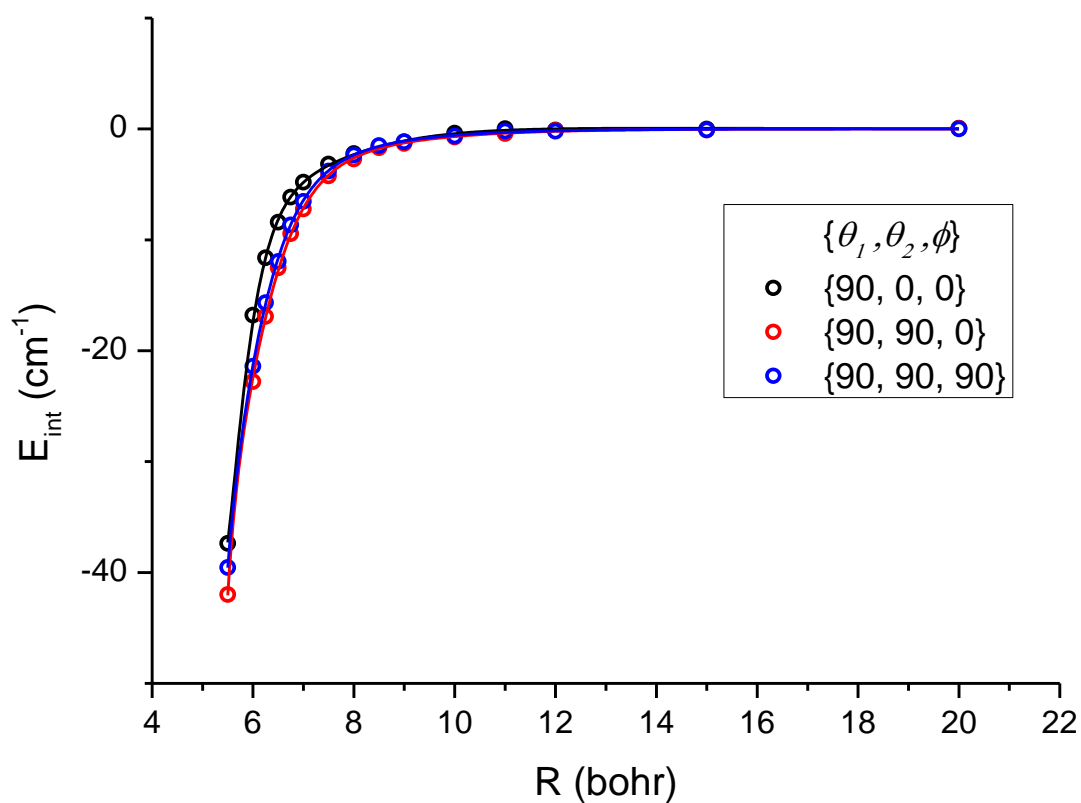
**Figure 4-7** Comparison between calculated interaction energies and fitted potential energy surfaces for $\Delta V$ along R. Dots in the figures were derived from the difference between averaged results (see 4.2.4 and 4.3) and CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.
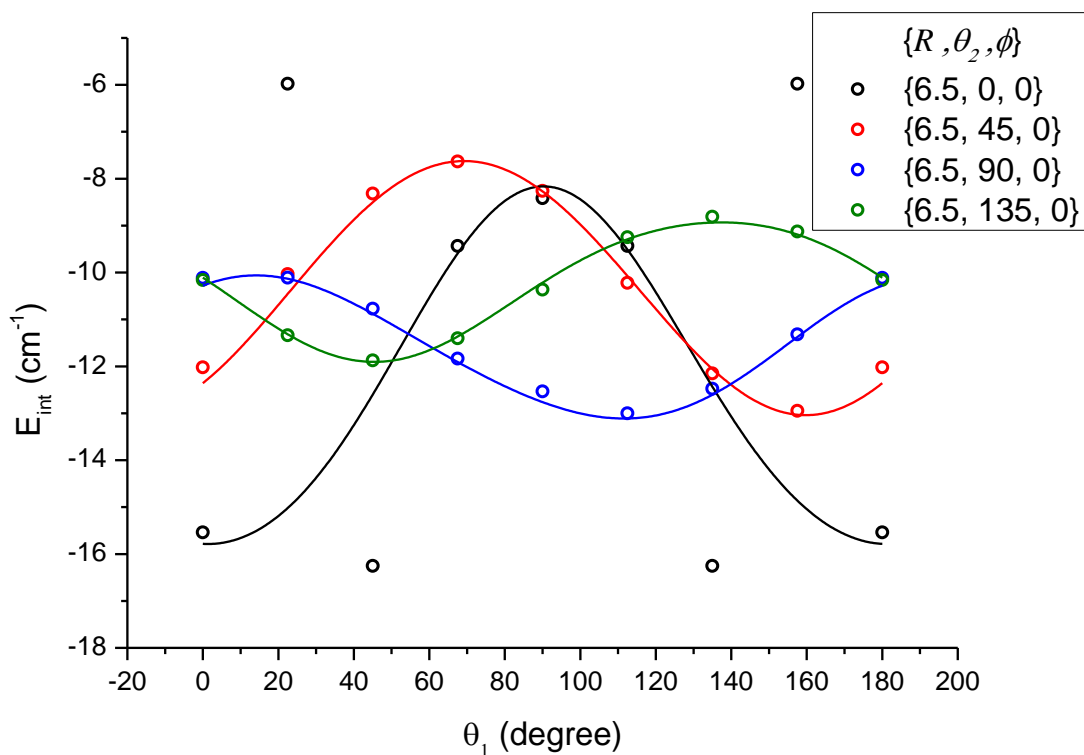
**Figure 4-8** Comparison between calculated interaction energies and fitted potential energy surfaces for $\Delta V$ along $\theta_1$. Dots in the figures were derived from the difference between averaged results (see 4.2.4 and 4.3) and CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions. Four values of $\{R, \theta_2, \phi\}=\{6.5, 0, 0\}$ at $\theta_1=22.5, 45, 135, 157.5$ are discarded.
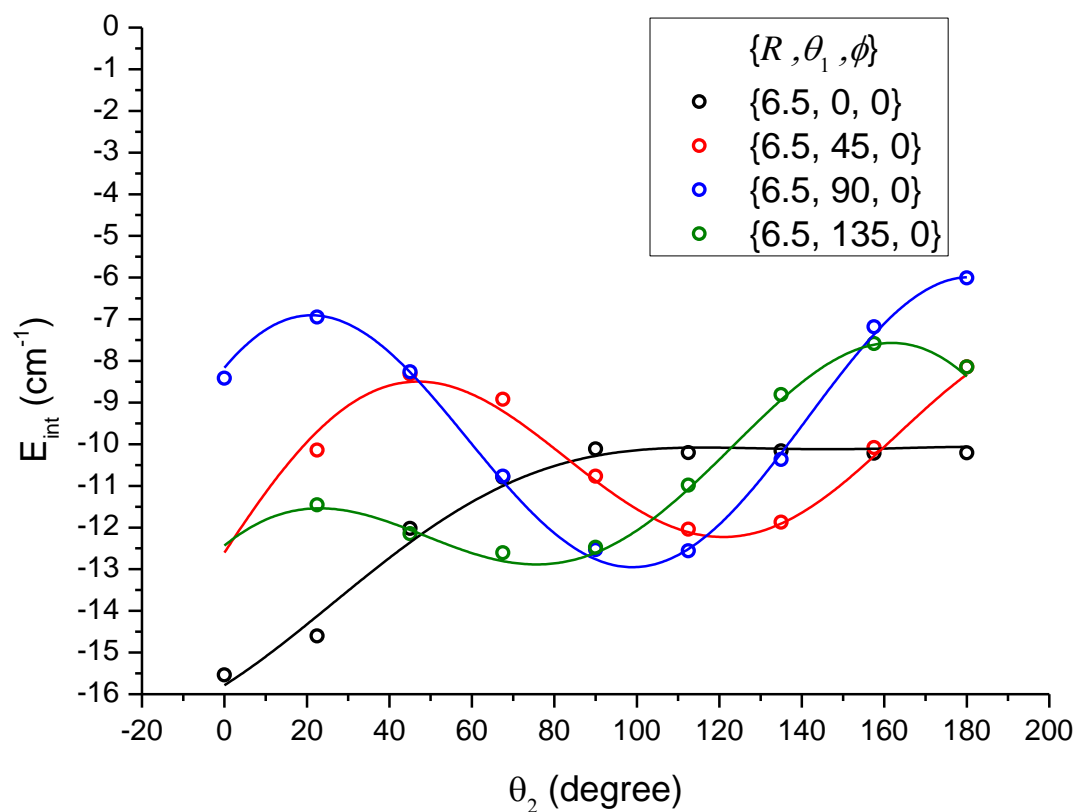
**Figure 4-9** Comparison between calculated interaction energies and fitted potential energy surfaces for $\Delta V$ along $\theta_2$. Dots in the figures were derived from the difference between averaged results (see 4.2.4 and 4.3) and CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.
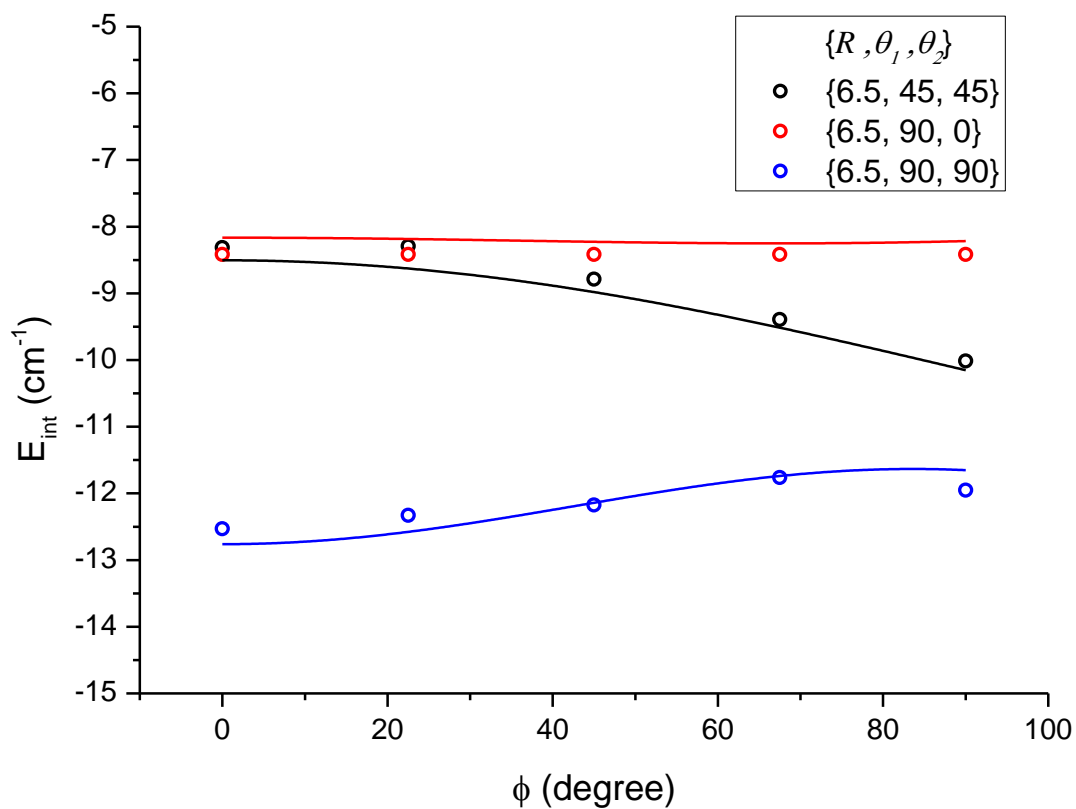
**Figure 4-10** Comparison between calculated interaction energies and fitted potential energy surfaces for $\Delta V$ along $\phi$. Dots in the figures were derived from the difference between averaged results (see 4.2.4 and 4.3) and CCSD(T)-F12b/AVTZ calculations. Lines were derived from the fitting functions.

**Table 4-20** The statistical information of the fitting errors for $\Delta V$.

| $\Delta V$ | |
| --- | --- |
| Errors (cm$^{-1}$) | All points in the grid[*] |
| Maximum | 3.352 |
| Mean | 0.084 |
| Standard deviation | 0.054 |



---

[*] Discarded points are not include.

### 4.4.3 Fitting results for $< V >_{00}$

Several geometries were chosen to illustrate the results of the fitting in $R$, $\theta_1$, $\theta_2$ and $\phi$ dimensions (Figure 4-11, Figure 4-12, Figure 4-13, Figure 4-14, respectively). The dots in the figures were derived from the averaged results (see 4.2.4 and 4.3). The lines were corresponding values from the fitting function $< V >_{00} = V_{F12b/AVTZ} + \Delta V$. The statistical information of the fitting error was shown in Table 4-21.
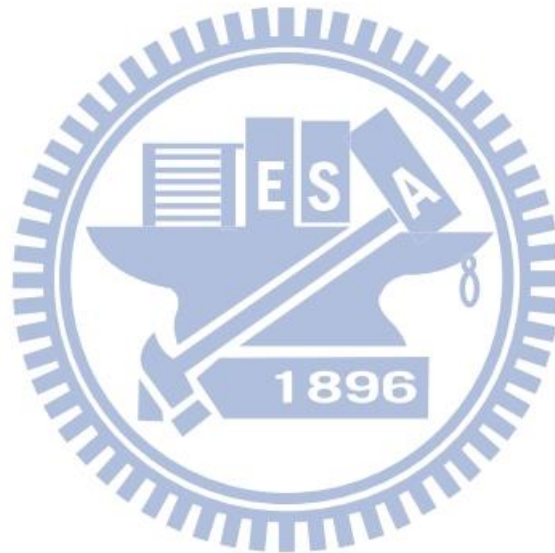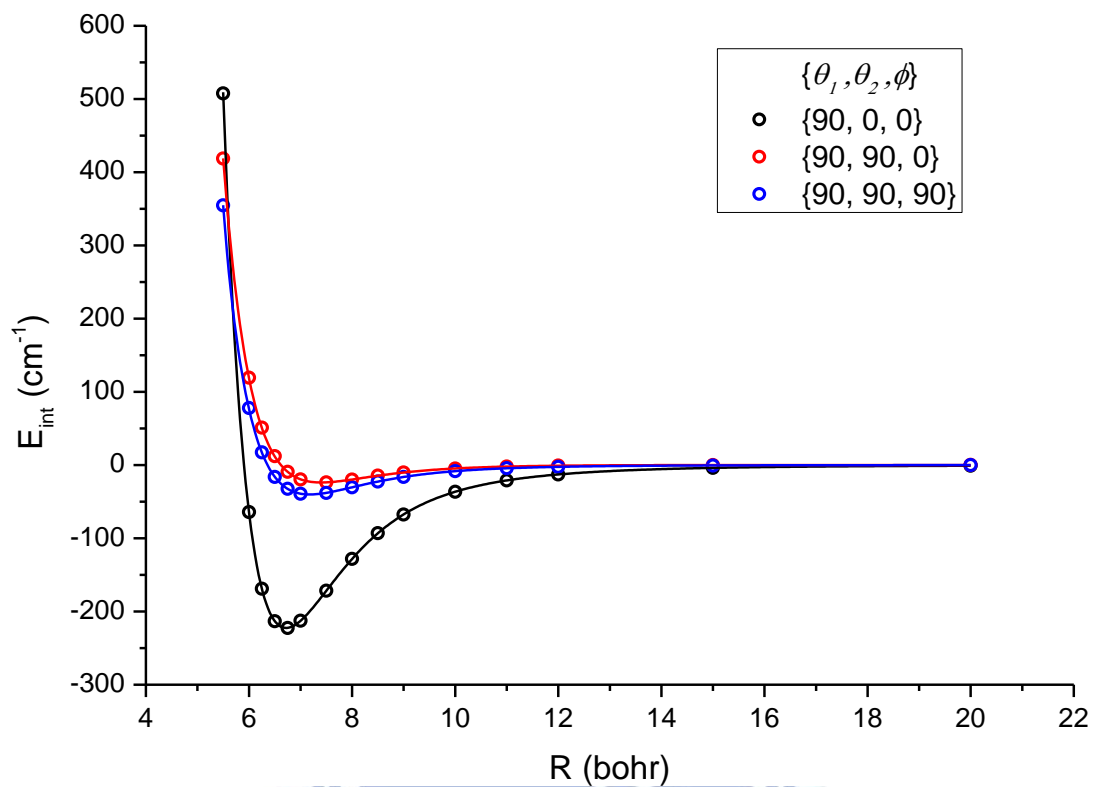
**Figure 4-11** Comparison between calculated interaction energies and fitted potential energy surfaces for $<V>_{00}$ along R. Dots in the figures were derived from the averaged results (see 4.2.4 and 4.3). Lines were derived from the resulting functions $<V>_{00} = V_{F12b/AVTZ} + \Delta V$.
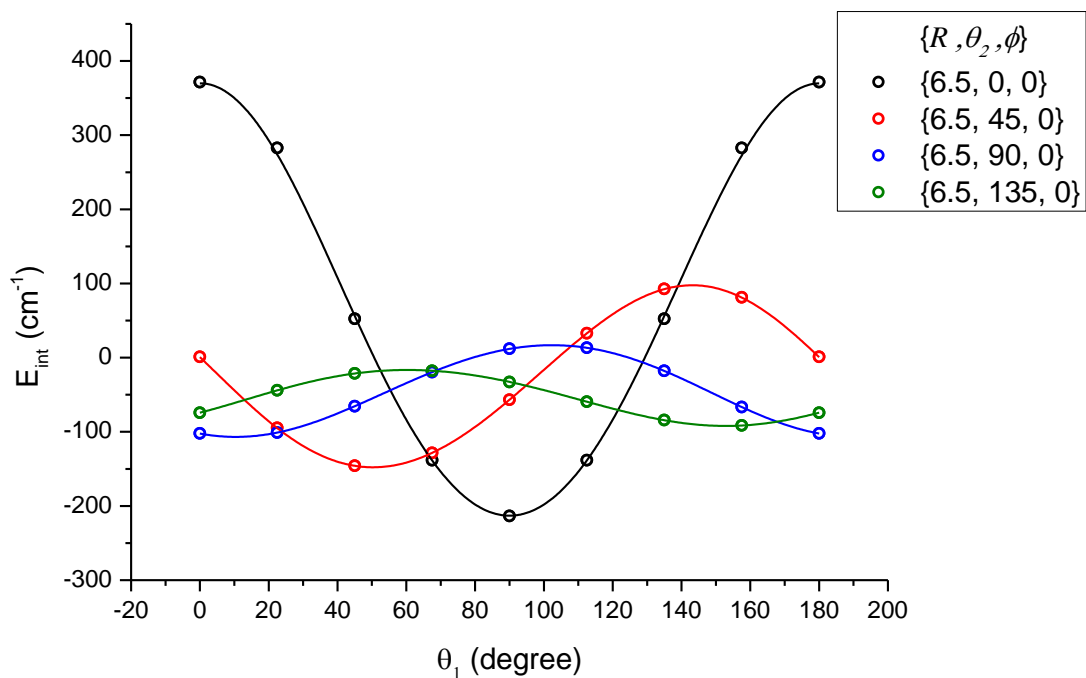
**Figure 4-12** Comparison between calculated interaction energies and fitted potential energy surfaces for $<V>_{00}$ along $\theta_1$. Dots in the figures were derived from the averaged results (see 4.2.4 and 4.3). Lines were derived from the fitting functions $<V>_{00} = V_{F12b/AVTZ} + \Delta V$.

**Figure 4-13** Comparison between calculated interaction energies and fitted potential energy surfaces for $<V>_{00}$ along $\theta_2$. Dots in the figures were derived from the averaged results (see 4.2.4 and 4.3). Lines were derived from the fitting functions $<V>_{00} = V_{F12b/AVTZ} + \Delta V$.

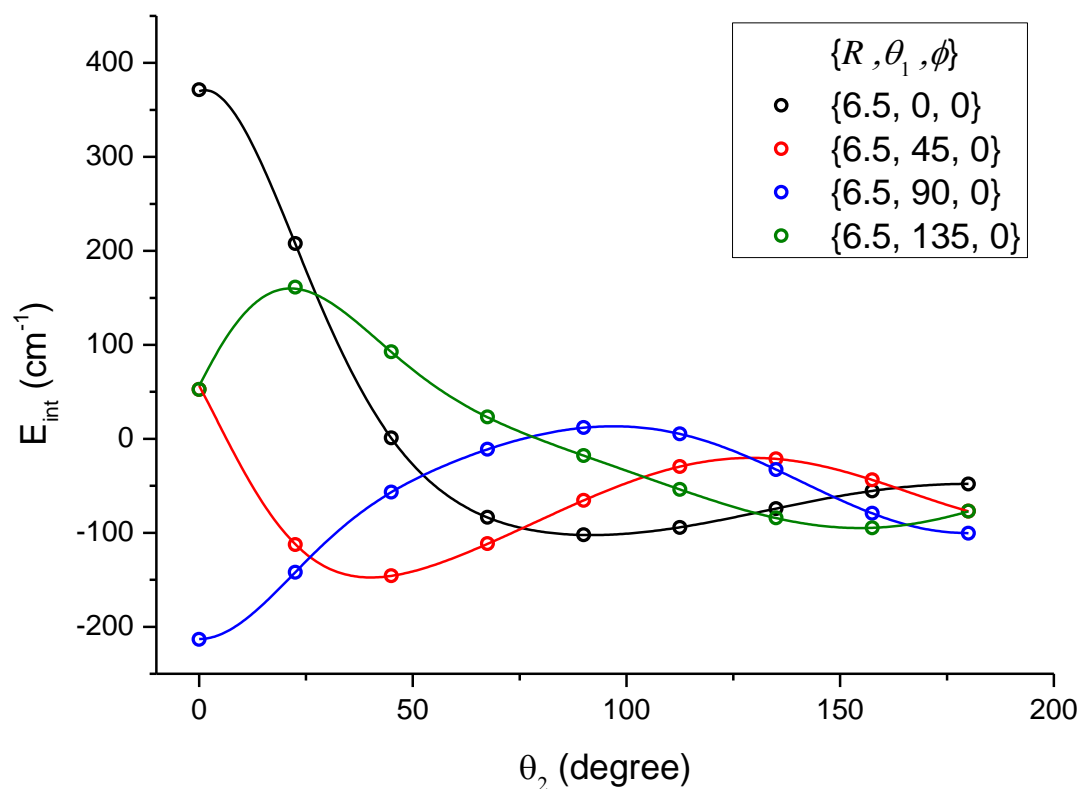**Figure 4-14** Comparison between calculated interaction energies and fitted potential energy surfaces for $< V >_{00}$ along $\phi$. Dots in the figures were derived from the averaged results (see 4.2.4 and 4.3). Lines were derived from the fitting functions $< V >_{00} = V_{F12b/AVTZ} + \Delta V$.

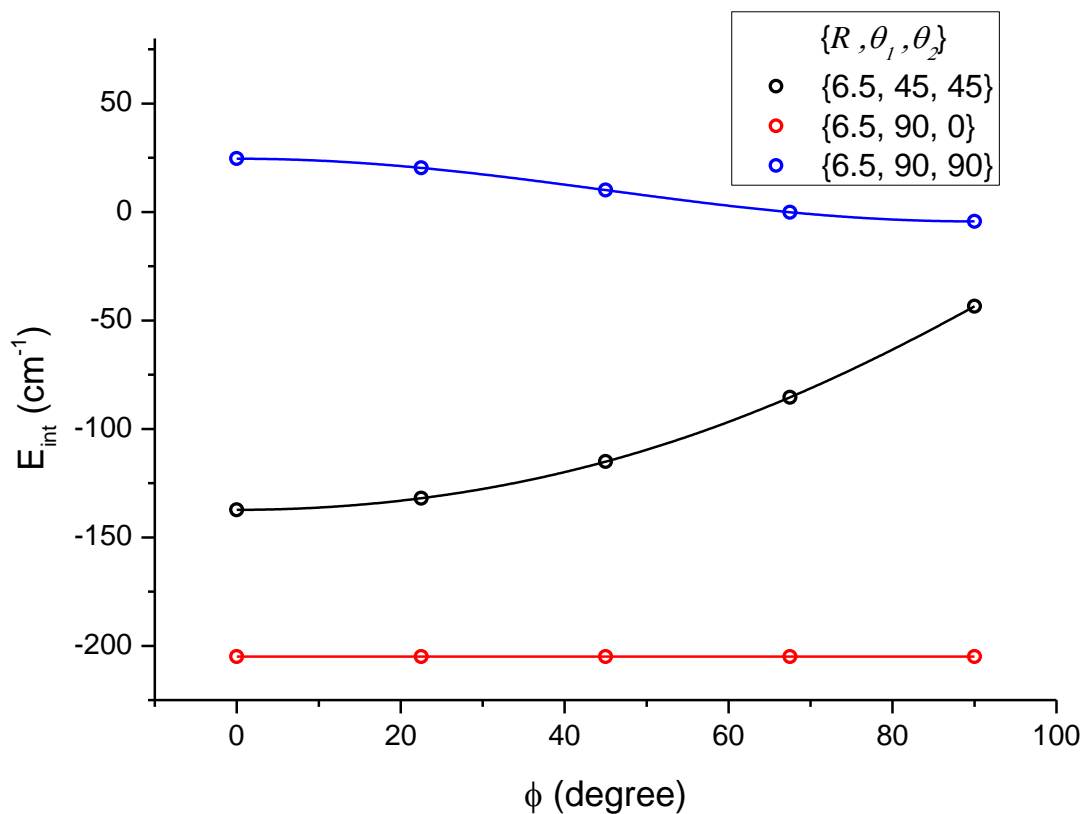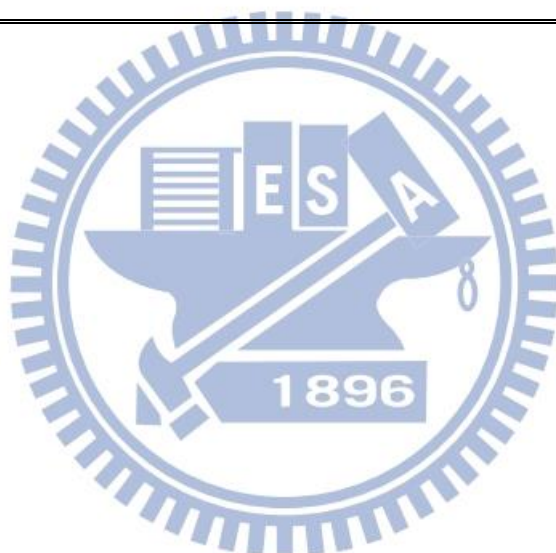**Table 4-21** The statistical information of the fitting errors for $< V >_{00}$.

| Errors ($cm^{-1}$) | $< V >_{00}$ | |
|---|---|---|
| | All points in the grid | Points with negative value of the interaction energy |
| Maximum | 10.17 | 0.47 |
| Mean | 0.192 | 0.066 |
| Standard deviation | 0.017 | 0.0021 |

## 4.4.4　Summary of the fit

From the results in this section, the two-step fitting has following benefits:

1.  The correct shape of the whole potential energy surface can be reproduced by denser and much faster CCSD(T)-F12b/AVTZ calculation.

2.  The correction $\Delta V$ from the fit of the difference between the most accurate results and CCSD(T)-F12b/AVTZ results can reproduce the interaction energy very well.

Also, the fitting functions are linear with respect to the fitting parameters makes the fitting procedure relatively simple. These simple trigonometric functions produced small fitting errors and are easy to manipulate in future work.
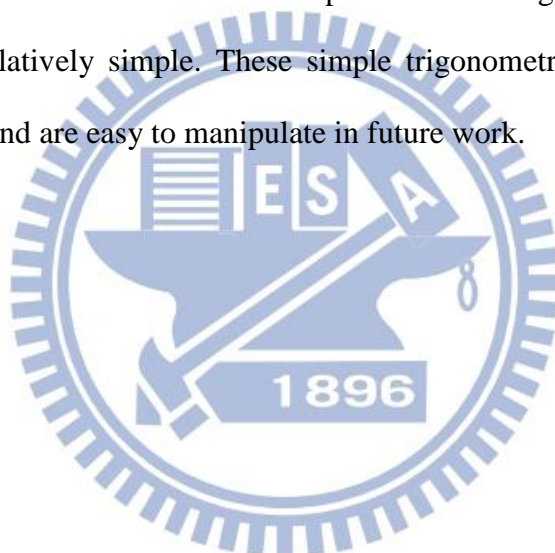
# Chapter 5

# Conclusion

A new four-dimensional interaction energy surface for the $H_2$–HCl complex is presented. The interaction energy values have been calculated with an averaging over the vibrational motion of $H_2$ and HCl.

The interaction energies at geometry near global minimum had been investigated thoroughly:

1. The interaction energy computed by CCSD(T)-F12b/CBS(AVQZ-AV5Z) approach could achieve comparable accuracy of standard CCSD(T) /CBS(AV5Z-AV6Z) with a strongly reduced computing cost.

2. The post-CCSD(T) correction $\delta E_{int}^{Q-(Q)}$ and higher coupled cluster excitations can be ignored entirely. The expensive accurate $\delta E_{int}^{(Q)-(T)}$ correction is handled by separating $\delta E_{int}^{(Q)-(T)}$ into $\delta E_{int}^{T-(T)}$ and $\delta E_{int}^{(Q)-T}$. The $\delta E_{int}^{T-(T)}$ decreases as the basis set size increases while the $\delta E_{int}^{(Q)-T}$ shows opposite trend. Therefore calculating $\delta E_{int}^{T-(T)}$ in a larger AVTZ basis set and $\delta E_{int}^{(Q)-T}$ in AVDZ set gave very good results with small uncertainty (0.3 cm$^{-1}$). The main benefit comes from opposite trends in the corrections to the interaction energy for CCSDT ($\delta E_{int}^{T-(T)}$) and CCSDT(Q) ($\delta E_{int}^{(Q)-T}$) approaches with increasing basis set size.

3. In core electron correction $\delta E_{int}^{AE}$ calculations, standard AV$X$Z sets converge slowly. ACV$X$Z and AWCV$X$Z results are consistent.

4. In relativistic correction $\delta E_{int}^{rel}$ calculations, most reasonable basis sets gave very close results except AV$X$Z sets. This might be due to the insufficient flexibility of the sets in the large-exponent range. Decontracted bases are a better choice. In this Thesis I choose dAVQZ basis set to avoid convergence problem in dACV$X$Z sets.

5. Midbond functions increase the convergence rate of the interaction energy $E_{int}^{CCSD(T)-F12b}$ and $E_{int}^{CCSD(T)}$. But they do not show improvement for $\delta E_{int}^{AE}$, $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{rel}$ calculations.

The adopted theory level and basis set are later tested for another 3 different geometries. All results shows uncertainty lower than 0.6 cm$^{-1}$. This indicates that the whole potential energy surfaces can be generated accurately by the method proposed here.

The interaction energies have been obtained as a sum of $E_{int}^{CCSD(T)-F12b}$, $\delta E_{int}^{AE}$, $\delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{rel}$ contributions calculated in different-size basis sets up to aug-cc-pV5Z for the grid points corresponding to the reference intramonomer separations. $E_{int}^{CCSD(T)-F12b}$ and $\delta E_{int}^{AE}$ results were extrapolated to the infinite basis set limit. For the energies used for calculations of Taylor-expansion derivatives, the fairly time-consuming $\delta E_{int}^{AE}, \delta E_{int}^{T-(T)}$, $\delta E_{int}^{(Q)-T}$, and $\delta E_{int}^{rel}$ terms could be neglected. The uncertainty caused by vibrationally averaged approximation can be ignored.

The two-step fitting was proposed in the Thesis which reproduced the correct shape of the whole potential energy surface and the interaction energy well. The maximum error of the fitting in the attractive part of potential is 0.47 cm$^{-1}$ and the mean error is 0.0066 cm$^{-1}$. The fitting functions are linear with respect to the fitting

parameters what makes the fitting procedure relatively simple. The simple trigonometric functions were used which are easy to manipulate in future work.

This excellent accuracy (0.6 cm$^{-1}$) is comparable to the similar *ab initio* H$_2$–CO surfaces of Jankowski *et al*[3] even in this larger system. The Thesis also demonstrated with the powerful computational power nowadays, that it is possible to generate accurate high dimensional interaction energy surfaces of many-electrons systems. Similar interaction energy surfaces (e.g. H$_2$–HX, X=F, Br) can be produced based on the methodology developed in the presented Thesis. The quality of the potential energy surfaces can be tested by comparing the second virial coefficient between experiment and calculated value. The potential will be used for simulating the spectrum of HCl trapped in pH$_2$ to understand more elementary mechanism in the field of atmosphere chemistry.
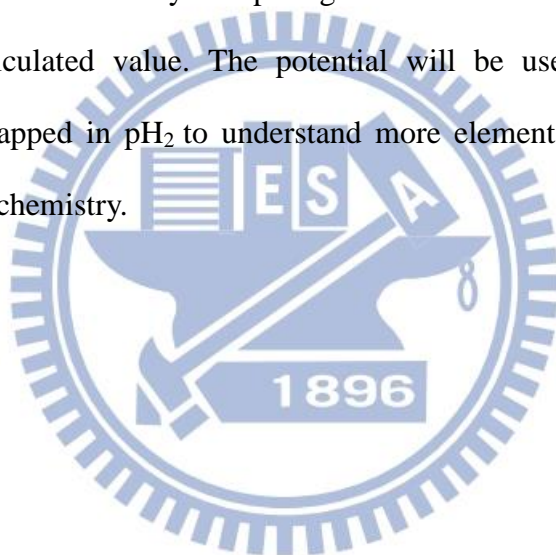
# References

[1] P. Jankowski, *Journal of Chemical Physics* **2004**, *121*, 1655-1662.

[2] P. Jankowski and K. Szalewicz, *Journal of Chemical Physics* **2005**, *123*.

[3] P. Jankowski, L. Surin, A. Potapov, S. Schlemmer, A. McKellar and K. Szalewicz, *The Journal of Chemical Physics* **2013**, *138*, 084307.

[4] D. T. Anderson, R. J. Hinde, S. Tam and M. E. Fajardo, *Journal of Chemical Physics* **2002**, *116*, 594-607.

[5] M. Born and J. R. Oppenheimer, *Annalen der Physik (Leipzig)* **1927**, *84*, 457-484.

[6] M. Born and K. Huang, *Dynamical theory of crystal lattices*, Clarendon Press, Oxford,, **1954**, p. xii, 420 p.

[7] a) B. Jeziorski, R. Moszynski and K. Szalewicz, *Chemical Reviews* **1994**, *94*, 1887-1930; b) G. Chalasinski and M. M. Szczesniak, *Chemical Reviews* **1994**, *94*, 1723-1765.

[8] J. C. Slater, *Quantum theory of molecules and solids*, McGraw-Hill, New York,, **1963**, p.

[9] I. G. Kaplan, *Intermolecular interactions : physical picture, computational methods, model potentials*, Wiley, Hoboken, NJ, **2006**, p. xii, 367 p.

[10] T. B. Adler, G. Knizia and H. J. Werner, *Journal of Chemical Physics* **2007**, *127*.

[11] K. Patkowski, *The Journal of Chemical Physics* **2013**, *138*, 154101.

[12] S. Ten-no, *Abstracts of Papers of the American Chemical Society* **2004**, *228*, U241-U241.

[13] G. Knizia, T. B. Adler and H. J. Werner, *Journal of Chemical Physics* **2009**, *130*.

[14] H. B. Jansen and P. Ros, *Chemical Physics Letters* **1969**, *3*.

[15] S. F. Boys and F. Bernardi, *Molecular Physics* **2002**, *100*, 65-73.

[16] B. Liu and A. D. Mclean, *Journal of Chemical Physics* **1973**, *59*.

[17] a) F. M. Tao and Y. K. Pan, *Journal of Chemical Physics* **1992**, *97*, 4989-4995; b) F. M. Tao, *The Journal of Chemical Physics* **1993**, *98*, 3049-3059.

[18] a) H. J. Werner and P. J. Knowles, *Journal of Chemical Physics* **1985**, *82*, 5053-5063; b) P. J. Knowles and H. J. Werner, *Chemical Physics Letters* **1985**, *115*, 259-267.

[19] M. Kallay and P. R. Surjan, *Journal of Chemical Physics* **2001**, *115*, 2945-2954.

[20] F. B. Vanduijneveldt, J. G. C. M. Vanduijneveldtvanderijdt and J. H. Vanlenthe, *Chemical Reviews* **1994**, *94*, 1873-1885.

[21] a) A. Wolf, M. Reiher and B. A. Hess, *Journal of Chemical Physics* **2002**, *117*, 9215-9226; b) M. Reiher and A. Wolf, *Journal of Chemical Physics* **2004**, *121*,

2037-2047; c) M. Reiher and A. Wolf, *Journal of Chemical Physics* **2004**, *121*, 10945-10956.

[22] T. H. Dunning, *Journal of Chemical Physics* **1989**, *90*, 1007-1023.

[23] T. H. Dunning, K. A. Peterson and A. K. Wilson, *Journal of Chemical Physics* **2001**, *114*, 9244-9253.

[24] K. A. Peterson and T. H. Dunning, *Journal of Chemical Physics* **2002**, *117*, 10548-10560.

[25] a) F. Weigend, A. Kohn and C. Hattig, *Journal of Chemical Physics* **2002**, *116*, 3175-3183; b) C. Hattig, *Physical Chemistry Chemical Physics* **2005**, *7*, 59-66.

[26] F. Weigend, *Physical Chemistry Chemical Physics* **2002**, *4*, 4285-4291.

[27] S. Ten-No, *Chemical Physics Letters* **2004**, *398*, 56-61.

[28] O. Marchetti and H. J. Werner, *Journal of Physical Chemistry A* **2009**, *113*, 11580-11585.

[29] a) A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chemical Physics Letters* **1998**, *286*, 243-252; b) T. Helgaker, W. Klopper, H. Koch and J. Noga, *Journal of Chemical Physics* **1997**, *106*, 9639-9646.

[30] a) W. Klopper and C. C. M. Samson, *Journal of Chemical Physics* **2002**, *116*, 6397-6410; b) W. Kutzelnigg and W. Klopper, *Journal of Chemical Physics* **1991**, *94*, 1985-2001.

[31] P. Jankowski.

[32] P. Huxley and J. N. Murrell, *Journal of the Chemical Society-Faraday Transactions Ii* **1983**, *79*, 323-328.

[33] J. Tennyson, S. Miller and C. R. Lesueur, *Computer Physics Communications* **1993**, *75*, 339-364.

[34] S. Bubin and L. Adamowicz, *Journal of Chemical Physics* **2003**, *118*, 3079-3082.

[35] M. Slawik, private communication, michal.slawik@gmail.com.