

ACCURATE POLARIZATION ENERGIES FROM FORCE FIELDS

by

Dongliang Yang

B.S., University of Science and Technology of China, 2008

Submitted to the Graduate Faculty of
Arts and Science in partial fulfillment
of the requirements for the degree of
Master of Science

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH
SCHOOL OF ARTS AND SCIENCE

This thesis was presented

by

Dongliang Yang

It was defended on

August 6th, 2010

and approved by

Dr. Kenneth Jordan, Professor, Department of Chemistry

Dr. Rob Coalson, Professor, Department of Chemistry

Dr. Geoffrey Hutchison, Assistant Professor, Department of Chemistry

Thesis Director: Dr. Kenneth Jordan, Professor, Department of Chemistry

Copyright © by Dongliang Yang

2010

ACCURATE POLARIZATION ENERGIES FROM FORCE FIELDS

Dongliang Yang, M.S.

University of Pittsburgh, 2010

The distributed point polarizable water model (DPP2) [Ref 1] which was recently developed in our group, has explicit terms for induction, charge penetration, and charge transfer. It is a refinement of the DPP model [Ref 2] which was also developed in our group. The DPP2 model has been found to accurately describe the interaction energy in water clusters. In this work, we aim to further improve the accuracy of DPP2 in calculating the induction energy. There are two ways to model higher-order polarization effects, one is through distributed atomic dipole polarizabilities, the other is through single-center expansion with higher multipoles. We developed a fitting method which can map the distributed dipole polarizabilities into the dipole-dipole (α), dipole-quadrupole (A) and quadrupole-quadrupole (C) polarizabilities. DPP2 uses three distributed atomic polarizabilities $\alpha_0, \alpha_{H1}, \alpha_{H2}$ which are located on oxygen and hydrogen atoms to describe induction effects. We show that the A and C values associated with the DPP2 model differ appreciably from the results of high level ab-initio methods [Ref 3, 4, 5, 25]. We have explored several strategies for improving on the DPP2 results. We describe a 4-site polarizable model, in which we split the polarizability of oxygen onto its two lone-pair sites, and which gives results in good agreement with the ab-initio calculations.

TABLE OF CONTENTS

LIST OF TABLES	VII
LIST OF FIGURES	VIII
ACKNOWLEDGMENTS	IX
1.0 INTRODUCTION.....	1
1.1 CLASSIFICATION OF INTERMOLECULAR FORCES.....	1
1.2 INTRODUCTION TO DPP2 MODEL.....	3
1.3 THE IMPORTANCE OF ACCURATE C VALUES.....	4
2.0 THEORY	5
2.1 MULTIPOLE MOMENTS.....	5
2.1.1 Cartesian tensor definition	6
2.1.2 Spherical tensor expression	7
2.1.3 Change of origin.....	8
2.2 POLARIZATION THEORY.....	8
2.2.1 Definition of polarizabilities from perturbation theory	8
2.2.2 The energy of a molecule in a non-uniform electric field	9
2.2.3 Classical treatment of polarizabilities.....	11
2.2.4 Symmetry in polarizabilities.....	12
2.2.5 Single-center expansion versus distributed polarizabilities.....	14

2.2.6	How to calculate polarizabilities in practice	15
3.0	INDUCTION ENERGY ASSOCIATED WITH DPP2 MODEL	18
4.0	DETAILS OF THE CALCULATIONS.....	22
4.1	HOW TO ESTIMATE THE C CONTRIBUTION:.....	22
4.2	FITING METHOD	23
4.3	EVALUATING DIPOLE-QUADRUPOLE POLARIZABILITIES	27
4.4	HOW TO VERIFY THIS METHOD.	29
5.0	RESULTS AND DISCUSSION	30
5.1	DPP AND DPP2	30
5.2	CHANGE THE GEOMETRY.....	31
5.3	CHANGE POLARAZIBILITY FROM OXYGEN TO M SITE	32
5.4	FOUR-SITE POLARIZABLE MODELS	33
6.0	CONCLUSION.....	36
	REFERENCES.....	37

LIST OF TABLES

Table 1 Contributions to the energy of interaction between molecules.....	2
Table 2 Character table for C_{2v} point group.....	12
Table 3 Product table for C_{2v} point group.....	13
Table 4 Geometry of the H_2O molecule (Angstrom), from Bishop	15
Table 5 Geometry of the H_2O molecule (Angstrom) in the DPP2 model	19
Table 6 Parameters in DPP and DPP2 model.....	21
Table 7 C values from ASP –W4 and our fitting method.....	29
Table 8 Values of C components of DPP1 and DPP2	31
Table 9 α and C values of polarizable M site DPP2 model	32
Table 10 Parameters used in seven lone-pair 4-site polarizable models.	34
Table 11 α and C values of various 4 site polarizable DPP2 model.....	35

LIST OF FIGURES

Figure 1 Energy contributions from α , A, and C polarizabilities of the water dimer as a function of the O-O distance, produced from ASP-W4 model using Orient software developed by Dr. Anthony Stone, etc.	4
Figure 2 DPP2 water with its parameters	18
Figure 3 Schematic structure of DPP2M model	32
Figure 4 Schematic structure of 4-site polarizable model	33

ACKNOWLEDGMENTS

First I wish to express my sincere appreciation to my research advisor Professor Kenneth Jordan. It is he who taught me what a scientific attitude should be like. To keep a self motivation, to keep thinking in a bigger picture but at the same time to be consider all the details is what I learned from him. His passion in science, his great wide and deep knowledge, and serious attitude toward science impress me a deeply. Whatever career I do in my future, I shall always keep a serious scholar's attitude and passion to my career and life like he does. Secondly, I want to give my thanks to Dr. Fangfang Wang. During this project, we worked on similar problem and had many discussions, where we helped each other and together moved the projects forward. Also I want to thank Dr. Revati Kumar, who developed the DPP2 model. I want to thank Fangyong Yan and Glen Jenness for their support. I also want to give my thanks to all the other group members for all the interesting scientific discussions, their encouragement and various suggestions. I want to give my thanks to my committee members. I want to thank them for their patience, their time and their help. Their great questions help me to give rethinking on my projects, at which angle I never thought before.

I also want to give my thanks to University of Pittsburgh. I learned a lot in two years study here. I still remember Dr. Asher's special attention to every student in his class. Classical Literature Reading in Neuroscience taught by Dr. Schwartz let me deeply appreciate the

greatness of scientific logical thinking and experiment designs. Also I miss my students in my TA labs a lot.

I also want to give my thanks to natural beauty in the city of Pittsburgh. It is a great city for academic and scientific research. I want to give my thanks to the great Schenley Park, where I could play soccer at every weekend there and meet a lot international friends, where I could walk on the little quiet path through hills leading my from Oakland to squirrel hill. I like Ohio State Park where I could do rafting. I like Seven Spring where I could play snowboarding; I like Moraine State Park where I really enjoy the beauty of lake and the colorful tress surrounding it. Also the snow in the last winter is the biggest one I ever met, which really shock me for how powerful the nature could be.

At last I want to give my special thanks to my marathon experience in the city of Pittsburg. It was my first time to run a full marathon. The experience told me that most people can arrive at much higher achievements if they just set the gold and go for it, no matter how smart they are, what background they have. With the reasonable great goal you really dreamed of in mind and run for it, don't care to much about difficulties on the road ahead, just go on and on, solve the problems we encountered on road one by one, never give up! After all, we won't regret the entire road we passed by where we used our heart to do it.

1.0 INTRODUCTION

When developing a force field, one needs to classify molecular interactions into different forces based on physical sense. This chapter will proceed as following: In section 1.1, I will introduce the traditional classification of intermolecular forces and their special features and describe the role the polarization energy play in a force field. In section 1.2, we provide a brief introduction to our DPP2 force field. In section 1.3, I will give the reasons why we are interested in developing a more accurate force field.

1.1 CLASSIFICATION OF INTERMOLECULAR FORCES

In classical models, intermolecular forces are classified as electrostatic energy, polarization energy, dispersion energy, repulsion energy, etc. These interactions can be separated into long-range and short range interactions. The energies of long-range interactions behave as some inverse power of R , and the energies of short-range interactions decrease exponentially with distance.

Long-range effects include electrostatic, induction and dispersion. Electrostatic effects arise from the straightforward classical interaction between the static charge distributions, which

are strictly pairwise additive. Induction effects arise from the distortion of charge distribution in the external electric field and are always attractive. Induction energies are non-additive, since they do not behave linearly on the electrical field created by neighbor charges. Dispersion effects arise from the correlated fluctuations of electrons on different sites.

At short ranges where the molecular wavefunctions overlap significantly, exchange and repulsion are most important, and often taken together and described as exchange-repulsion. Penetration, charge transfer and damping are modifications of the long-range terms arising from the overlap of the wavefunctions. Charge transfer interactions are often included as a part of the induction energy.

Table 1 Contributions to the energy of interaction between molecules

Contribution	Additive?	sign	Comments
Long-range($U \sim R^{-n}$)			
Electrostatic	Yes	\pm	Strong orientation dependence
Induction	No	-	
Dispersion	Approx	-	Always present
Short-range($U \sim e^{-\alpha R}$)			
Exchange	No	-	
Repulsion	No	+	Dominates at very short range
Charge Transfer	No	-	Donor-acceptor interaction
Penetration	Yes	-	Can be repulsive at very short range
Damping	Approx	+	Modification of dispersion and induction

Note: This table is reproduced from A.J. Stone's Book "The Theory of Intermolecular Forces"
[Ref 6].

The relationship between induction energy and polarization energy is: In quantum chemistry, the total induction energy embodies several terms: polarization, nonclassical charge-transfer, and exchange-induction energies, each with different physical origin [Ref 7-10]. A simple polarizable force field using distributed dipole polarizability approach can deal with the polarization part of the induction.

1.2 INTRODUCTION TO DPP2 MODEL

The DPP2 water model was recently developed in our group [Ref 1]. It has explicit terms for polarization, charge penetration, and charge transfer, besides electrostatics, dispersion, and exchange-repulsion. One of our purposes of developing this model is to accurately calculate the interaction energies of water clusters and to describe their vibrational spectroscopy, which is quite sensitive to polarization energy because of its dependence on the second derivatives of the energy. Also, when considering the non-additivity of polarization interactions, the larger the water cluster is, the greater the contribution the polarization energy will make to the total interaction energy. So in order to describe the vibrational spectroscopy properties of water, an accurate calculation of polarization energy is crucial.

1.3 THE IMPORTANCE OF ACCURATE C VALUES

“C” is the symbol for quadrupole-quadrupole polarizability. Similarly “A” is the symbol for dipole-quadrupole polarizability and “ α ” is the symbol for dipole-dipole polarizability. We will explain them in detail in Chapter 2. Calculations using the ASP-W4 water model [Ref 11] and Orient Software develop by Dr. Anthony Stone, etc. [Ref 32] show that α , A, C contribute -2.17, -0.28, and -0.77 kJ/mol respectively of a total induction energy of -3.22 kJ/mol for water dimer at the equilibrium geometry [Figure1]. The details of the calculations will be described in Section 4.1. As a result we cannot neglect the C terms if we want to accurately calculate the induction energy.

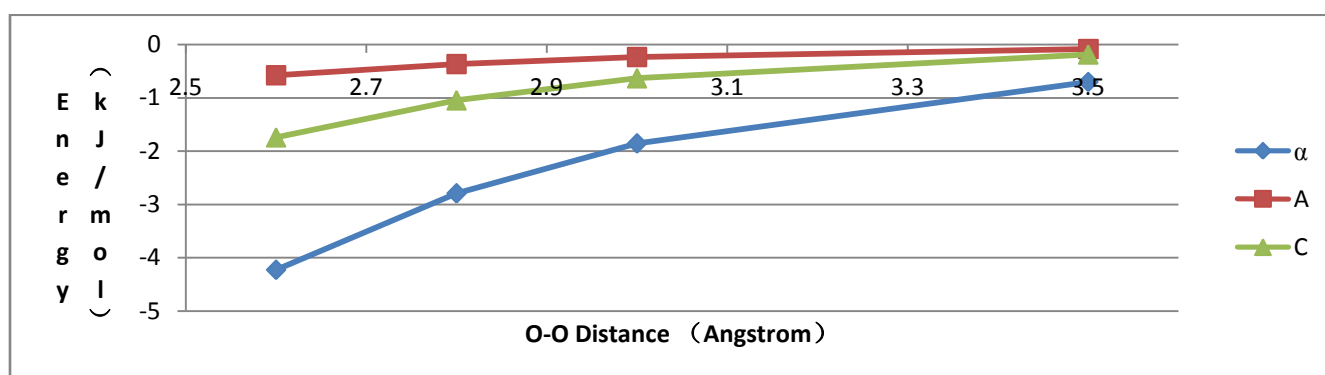


Figure 1 Energy contributions from α , A, and C polarizabilities of the water dimer as a function of the O-O distance, produced from ASP-W4 [Ref 11] model using Orient software developed by Dr. Anthony Stone, etc. [Ref 32].

2.0 THEORY

Multipole moments are used to describe the way in which the charge is distributed in a molecule. The theory of multipole moments and its two mathematical expressions will be introduced in Section 2.1.1 and 2.1.2. There have been different multipole moments values reported for the same system, which has caused a lot of confusion. Section 2.1.3. shows that different definitions of origin lead to different values of multipole moments. The theory of polarization will be the major focus of this chapter. I will first give the definitions of polarizabilities from a quantum mechanics view. Then in Section 2.2.2 and 2.2.3 I will describe how the polarization energy is treated in a classical way. Taylor series expansions will be used here. The quadrupole-quadrupole polarizability could have up to 81 different components, but in general the number of unique non-zero component is much smaller. This will be shown in section 2.2.4. In section 2.2.5, I will talk about the advantages and disadvantages of single versus multicenter expansions. In the last section, I will introduce two practical methods to calculate the higher rank polarizabilities of water.

2.1 MULTIPOLE MOMENTS

This sections includes subsection 2.1.1, 2.1.2, and 2.1.3.

2.1.1 Cartesian tensor definition

The most familiar multipole moment is the total charge: $q = \int \rho(\mathbf{r})d\mathbf{r}$, where $\rho(\mathbf{r})$ is the charge density at position \mathbf{r} . Since we will define dipole moments and other higher moments later, we do the sum of point charges here instead of integrating over electron distributions, $q = \sum_a e_a$, where e_a is the point charge on particle a and the sum is taken over all the electrons and nuclei.

The next one is the dipole moment:

$$\hat{\mu}_\alpha = \sum_a e_a a_\alpha \quad (2.1)$$

where e_a is the charge on particle a, a_α is the position of particle a and α stands for x, y, z.

The definition of the quadrupole moment components are

$$\hat{\Theta}_{zz} = \sum_a e_a a^2 \left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right) \quad (2.2)$$

$$\hat{\Theta}_{xx} = \sum_a e_a \left(\frac{3}{2} a_x^2 - \frac{1}{2} a^2 \right) \quad (2.3)$$

$$\hat{\Theta}_{yy} = \sum_a e_a \left(\frac{3}{2} a_y^2 - \frac{1}{2} a^2 \right) \quad (2.4)$$

$$\hat{\Theta}_{xy} = \frac{3}{2} \sum_a e_a a_x a_y \quad (2.5)$$

$$\hat{\Theta}_{xz} = \frac{3}{2} \sum_a e_a a_x a_z \quad (2.6)$$

$$\hat{\Theta}_{yz} = \frac{3}{2} \sum_a e_a a_y a_z \quad (2.7)$$

Notice that

$$\hat{\Theta}_{xx} + \hat{\Theta}_{yy} + \hat{\Theta}_{zz} = 0 \quad (2.8)$$

as a direct consequence of the definition. If we use the Einstein summation, the expression for quadrupole moments are:

$$\hat{\Theta}_{\alpha\beta} = \sum_a e_a \left(\frac{3}{2} a_\alpha a_\beta - \frac{1}{2} a^2 \delta_{\alpha\beta} \right) \quad (2.9)$$

2.1.2 Spherical tensor expression

We can also use spherical tensors to express multipole moments, which in some applications are more convenient to use.

$$\hat{Q}_{l\kappa} = \sum_a e_a R_{l\kappa}(\mathbf{a}) \quad (2.10)$$

or

$$Q_{l\kappa} = \int \rho(\mathbf{r}) R_{l\kappa}(\mathbf{r}) d^3\mathbf{r} \quad (2.11)$$

Here $\rho(\mathbf{r})$ stands for the charge density at position \mathbf{r} and $R_{l\kappa}(\mathbf{r})$ is the regular spherical harmonics. We use the label κ to denote a member of the series $0, 1c, 1s, 2c, 2s, \dots$. Where there is a sum over κ in a quantity labeled by $l\kappa$, the sum runs over the values $0, 1c, 1s, \dots, 1c, 1s$. The definitions of spherical tensors can be found in appendix B in Stone's book [Ref 6]. The conversion formulas between cartesian and spherical tensors can be found in appendix E in Stone's book [Ref 6].

2.1.3 Change of origin

The values of multipole moments depend on our choice of origin when their lower rank multipole moments are non-zero. In many applications it is convenient to take the origin at the center of mass, but it is not necessarily the optimum choice for describing the electrostatic properties. For the example of water, taking the oxygen as the origin is another popular choice. McLean derived the conversion formulas for all μ , Θ , Ω , α , A , C , etc. in his paper published in 1967[Ref12].

2.2 POLARIZATION THEORY

This section includes 6 subsections.

2.2.1 Definition of polarizabilities from perturbation theory

By perturbation theory, the definitions of polarizabilities are (details of those formulas can be found in Stone's book⁶ page 21):

$$\alpha_{\alpha\beta} = \sum_n' \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|0\rangle + \langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|0\rangle}{W_n - W_0} \quad (2.12)$$

$$A_{\alpha,\beta\gamma} = \sum_n' \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\Theta}_{\beta\gamma}|0\rangle + \langle 0|\hat{\Theta}_{\beta\gamma}|n\rangle\langle n|\hat{\mu}_\alpha|0\rangle}{W_n - W_0} \quad (2.13)$$

$$C_{\alpha\beta,\gamma\delta} = \frac{1}{3} \sum_n' \frac{\langle 0|\hat{\Theta}_{\alpha\beta}|n\rangle\langle n|\hat{\Theta}_{\gamma\delta}|0\rangle + \langle 0|\hat{\Theta}_{\gamma\delta}|n\rangle\langle n|\hat{\Theta}_{\alpha\beta}|0\rangle}{W_n - W_0} \quad (2.14)$$

“n” stands for the excited state n, “0” stands for the ground state, W_n is the eigenvalue of state n and W_0 is the eigenvalue of the ground state.

2.2.2 The energy of a molecule in a non-uniform electric field

Consider a molecule in an external potential $V(\mathbf{r})$. The electric field is $F_\alpha = -\frac{\partial V}{\partial r_\alpha} = -\nabla_\alpha V$. For a non-uniform electric field there is a field gradient $F_{\alpha\beta} = -\frac{\partial^2 V}{\partial r_\alpha \partial r_\beta} = -\nabla_\alpha \nabla_\beta V$. We can expand the potential in a Taylor series about a suitable origin and set of coordinate axes.

$$V(\mathbf{r}) = V(0) + r_\alpha V_\alpha(0) + \frac{1}{2} r_\alpha r_\beta V_{\alpha\beta}(0) + \frac{1}{3!} r_\alpha r_\beta r_\gamma V_{\alpha\beta\gamma}(0) + \dots \quad (2.15)$$

Here we are using the Einstein summation convention where a repeated suffix implies summation over the three axes x, y and z. The operator describing the energy of a molecule in the presence of this potential is

$$\mathcal{H}' = \sum_a e_a \hat{V}(\mathbf{a}) \quad (2.16)$$

Where the sum is taken over all the nuclei and electrons in the molecules; particle a is at position \mathbf{a} carrying charge e_a . Then

$$\mathcal{H}' = V(0) \sum_a e_a + V_a(0) \sum_a e_a a_\alpha + \frac{1}{2} V_{\alpha\beta} \sum_a e_a a_\alpha a_\beta + \dots \quad (2.17)$$

which we write as

$$\mathcal{H}' = \widehat{M}V + \widehat{M}_\alpha V_\alpha + \frac{1}{2}V_{\alpha\beta}\widehat{M}_{\alpha\beta} \quad (2.18)$$

Here we abbreviate $V_\alpha(0)$ to V_α , etc. and introduce the zeroth moment M , the first moment M_α , the second moment $M_{\alpha\beta}$, and so on. We can immediately identify the zeroth moment $\widehat{M} = \sum_\alpha e_\alpha$ with total charge q , and the first moment $\widehat{M} = \sum_\alpha e_\alpha \alpha$ with the dipole moment $\hat{\mu}_\alpha$.

The second moments are a little more complicated. We are interested only in the energy of the interaction with the field (equation 2.18). We define a new quantity $\widehat{M}'_{\alpha\beta} = \widehat{M}_{\alpha\beta} - k\delta_{\alpha\beta}$, where k is a constant and $\delta_{\alpha\beta}$ is the Kronecker tensor. Then

$$\begin{aligned} \frac{1}{2}V_{\alpha\beta}\widehat{M}'_{\alpha\beta} &= \frac{1}{2}V_{\alpha\beta}\widehat{M}_{\alpha\beta} - \frac{1}{2}k\delta_{\alpha\beta}V_{\alpha\beta} \\ &= \frac{1}{2}V_{\alpha\beta}\widehat{M}_{\alpha\beta} - \frac{1}{2}kV_{\alpha\alpha} \\ &= \frac{1}{2}V_{\alpha\beta}\widehat{M}_{\alpha\beta} \end{aligned} \quad (2.19)$$

where the last line follows from Laplace's equation:

$$V_{\alpha\alpha} = \nabla^2 V = 0 \quad (2.20)$$

This is true for any value of k . We now choose k so that $\widehat{M}'_{\alpha\beta}$ becomes traceless: $\widehat{M}'_{\alpha\alpha} \equiv \widehat{M}'_{xx} + \widehat{M}'_{yy} + \widehat{M}'_{zz} = 0$. Then $\widehat{M}'_{\alpha\alpha} - k\delta_{\alpha\alpha} = 0$, or $k = \frac{1}{3}\widehat{M}_{\alpha\alpha} = \frac{1}{3}\sum_\alpha e_\alpha a^2$. (Remember that $\delta_{\alpha\alpha} = 3$.)

Then we have

$$\widehat{M}'_{\alpha\beta} = \sum_a e_a \left(a_\alpha a_\beta - \frac{1}{3}a^2 \delta_{\alpha\beta} \right) = \frac{2}{3}\widehat{\Theta}_{\alpha\beta} \quad (2.21)$$

So by subtracting away the trace of $\widehat{M}_{\alpha\beta}$, which does not contribute to the electrostatic energy, we arrive at the quadrupole moment in the form given previously, except for a numerical factor.

The higher moments are manipulated in a similar way. When we modify $\widehat{M}_{\alpha\beta\gamma}$ so as to remove the trace terms that do not contribute to the electrostatic energy, we arrive at the octpole moment $\widehat{\Omega}_{\alpha\beta\gamma}$ and so on. The operator describing the interaction becomes

$$\mathcal{H}' = qV + \widehat{\mu}_\alpha V_\alpha + \frac{1}{3}\Theta_{\alpha\beta} V_{\alpha\beta} + \frac{1}{15}\Omega_{\alpha\beta\gamma} V_{\alpha\beta\gamma} + \dots \quad (2.22)$$

2.2.3 Classical treatment of polarizabilities

The energy of a molecule in a static electric field is given by Buckingham[Ref13]

$$\begin{aligned} W &= W^0 + W' + W'' + \dots \\ &= W^0 \\ &\quad + \mu_\alpha V_\alpha + \frac{1}{3}\Theta_{\alpha\beta} V_{\alpha\beta} + \frac{1}{15}\Omega_{\alpha\beta\gamma} V_{\alpha\beta\gamma} + \dots \\ &\quad - \frac{1}{2}\alpha_{\alpha\beta} V_\alpha V_\beta - \frac{1}{3}A_{\alpha,\beta\gamma} V_\alpha V_{\beta\gamma} - \frac{1}{6}C_{\alpha\beta,\gamma\delta} V_{\alpha\beta} V_{\gamma\delta} - \dots \end{aligned} \quad (2.23)$$

The derivative of the energy with respect to field gives:

$$\frac{\partial W}{\partial V_\xi} = \mu_\xi - \alpha_{\xi\beta} V_\beta - \frac{1}{3}A_{\xi,\beta\gamma} V_{\beta\gamma} - \dots \quad (2.24)$$

$$\mu_\xi^p = \mu_\xi + \alpha_{\xi\beta} F_\beta + \frac{1}{3}A_{\xi,\beta\gamma} F_{\beta\gamma} + \dots \quad (2.25)$$

“p” here means “perturbed”. So when $(\frac{\partial W}{\partial V_\xi})_v \rightarrow 0$, we get the static dipole moment μ_ξ (equation 2.25), and indeed this is commonly used as its definition, as well as the basis for some methods of calculating it. From a physical sense, we can see that $\alpha_{\xi\beta}$ describes the additional dipole

induced by an applied electric field F_β , and $A_{\xi,\beta\gamma}$ describes the dipole induced by an applied field gradient $F_{\beta\gamma}$.

Similarly,

$$3 \frac{\partial W}{\partial W_{\xi\eta}} = \Theta_{\xi\eta} - A_{\alpha,\xi\eta} V_\alpha - C_{\alpha\beta,\xi\eta} V_{\alpha\beta} - \dots \quad (2.26)$$

$$\Theta_{\xi\eta}^p = \Theta_{\xi\eta} + A_{\alpha,\xi\eta} F_\alpha + C_{\alpha\beta,\xi\eta} F_{\alpha\beta} + \dots \quad (2.27)$$

Equation 2.27 can be used to define the quadrupole moment. A also describes the quadrupole induced by an electric field, and C describes the quadrupole induced by a field gradient.

2.2.4 Symmetry in polarizabilities

Table 2 Character table for C_{2v} point group

	E	$C_2(z)$	$\sigma_v(xz)$	$\sigma_v(yz)$	linear,	Quadratic
A_1	1	1	1	1	Z	x^2, y^2, z^2
A_2	1	1	-1	-1	R_z	xy
B_1	1	-1	1	-1	x, R_y	xz
B_2	1	-1	-1	1	y, R_x	yz

We can use standard group-theoretical methods to discover the number of non-zero components of a multipole moment of a given rank, or to determine whether a particular moment is non-zero. All the multipole moments and polarizabilities must vanish unless they are totally symmetric.

Table 3 Product table for C_{2v} point group

	A_1	A_2	B_1	B_2
A_1	A_1	A_2	B_1	B_2
A_2	A_2	A_1	B_2	B_1
B_1	B_1	B_2	A_1	A_2
B_2	B_2	B_1	A_2	A_1

There are 27 A and 81 C terms for water. A simple way to see which terms are non-zero is to use the direct product table for the C_{2v} point group. Take α_{xx} for example. Since x direction has B_1 symmetry, the symmetry character of α_{xx} is $B_1 * B_1 = A_1$, which is totally symmetric. So α_{xx} is non zero. But the symmetry character of $A_{x,xx}$ is $B_1 * A_1 = B_1$ which is not totally symmetric. So $A_{x,xx}$ is zero.

In this moment we establish that there are eight non-zero A components left:

$A_{x:zx}, A_{x:zx}, A_{y:zy}, A_{y:yz}; A_{z:xx}, A_{z:xx}, A_{z:yy}, A_{z:zz}$

Among them, based on symmetry, we have the following relationship:

$$A_{x:xz} = A_{x:zx}, \quad A_{y:zy} = A_{y:yz}, \quad (2.28)$$

From definitions, we have the following relationship:

$$A_{z:xx} + A_{z:yy} + A_{z:zz} = 0 \quad (2.29)$$

There are 21 non-zero C components:

$C_{xx:xx}, C_{xx:yy}, C_{xx:zz}, C_{yy:xx}, C_{yy:zz}, C_{yy:zz}, C_{zz:xx}, C_{zz:yy}, C_{zz:zz},$

$C_{xy:xy}, C_{xy:yx}, C_{yx:xy}, C_{yx:yx}, C_{xz:xz}, C_{xz:zx}, C_{zx:xz}, C_{zx:zx}, C_{yz:yz}, C_{yz:zy},$
 $C_{zy:yz}, C_{zy:zy}$

Among them, based on symmetry there are following relationships:

$$\begin{aligned}
 C_{xx:yy} &= C_{yy:xx}, C_{xx:zz} = C_{zz:xx}, C_{yy:zz} = C_{zz:yy} \\
 C_{xy:xy} &= C_{xy:yx} = C_{yx:xy} = C_{yx:yx}, \\
 C_{xz:xz} &= C_{xz:zx} = C_{zx:xz} = C_{zx:zx} \\
 C_{yz:yz} &= C_{yz:zy} = C_{zy:yz} = C_{zy:zy}
 \end{aligned} \tag{2.30}$$

In addition:

$$\begin{aligned}
 C_{xx:xx} + C_{xx:yy}(C_{yy:xx}) + C_{xx:zz}(C_{zz:xx}) &= 0 \\
 C_{yy:xx}(C_{xx:yy}) + C_{yy:yy} + C_{yy:zz}(C_{zz:yy}) &= 0 \\
 C_{zz:xx}(C_{xx:zz}) + C_{zz:yy}(C_{yy:zz}) + C_{zz:zz} &= 0
 \end{aligned} \tag{2.31}$$

Those relationships will be used in our calculations later in Chapter 4.

2.2.5 Single-center expansion versus distributed polarizabilities

The polarizabilities describe changes in the charge redistribution perturbed by external fields. A single-center multipole description of those charge changes is subject to the convergence problems, which means that higher rank multipole moments may not be negligible. A distributed treatment is expected to give better result here, especially when the molecule is large and the description of charges in local regions is important.

A further consideration is the response of the molecular charge distribution to external fields from other molecules. Such fields are non-uniform, so that the strength of the field varies considerably from one part of the molecule to another. Taylor series describing the variation of

the field across the molecule converges poorly or not at all when the molecule is large. Here a distributed treatment automatically takes account of variations in the strength of the field, since we use the value of the field at each site rather than the value at some arbitrary origin, and the sphere of convergence around each site only has to extend far enough to enclose the region belonging to the site.

The distributed polarizabilities are not limited to atoms. It is also possible to attach meanings to bond or lone-pair polarizabilities.

2.2.6 How to calculate polarizabilities in practice

Besides using perturbation theory to calculate the polarizabilities, there are also other methods to calculate them [Ref 3, 14]. Here we will give a brief introduction of Bishop's method whose C values will be used as benchmark in our work.

Table 4 Geometry of the H₂O molecule (Angstrom), from Bishop [Ref 3].

Atom	x	y	z
O	0.00000000	0.00000000	-0.0656945
H1	-0.75753705	0.00000000	0.5213831
H2	0.75753705	0.00000000	0.5213831

Bishop placed a point charge Q or $-Q$ at R distance away from the center of mass of water and at different orientations. Then he uses MC SCF (multiconfigurational self consistent

field) calculations to get the values of perturbed dipoles and quadrupoles. Then he used equations such as the following (all the equations can be derived from equations 2.25 and 2.27 above):

$$C_{xx,xx} = \left[\Theta_{xx} \left(-Q, \frac{\pi}{2}, 0 \right) - \Theta_{xx} \left(Q, \frac{\pi}{2}, 0 \right) \right] \left(\frac{R^3}{6Q} \right) \quad (2.32)$$

$$C_{yy,yy} = \left[\Theta_{xx}(-Q, 0, 0) + \Theta_{xx}(-Q, \pi, 0) - \Theta_{xx}(Q, 0, 0) - \Theta_{xx}(Q, \pi, 0) \right. \\ \left. + \frac{6Q}{R^3} (C_{xx,xx} + C_{zz,zz}) \right] \left(\frac{R^3}{6Q} \right) \quad (2.33)$$

$$C_{zz,zz} = \left[\Theta_{zz}(-Q, 0, 0) + \Theta_{zz}(-Q, \pi, 0) - \Theta_{zz}(Q, 0, 0) \right. \\ \left. - \Theta_{zz}(Q, \pi, 0) \right] \left(\frac{R^3}{12Q} \right) \quad (2.34)$$

$$C_{xy,xy} = \left[\Theta_{xy} \left(-Q, \frac{\pi}{2}, \frac{\pi}{4} \right) - \Theta_{xy} \left(Q, \frac{\pi}{2}, \frac{\pi}{4} \right) \right] \left(\frac{R^3}{6Q} \right) \quad (2.35)$$

$$C_{xz,xz} = \left[\Theta_{xz} \left(-Q, \frac{\pi}{4}, 0 \right) - \Theta_{xz} \left(Q, \frac{\pi}{4}, 0 \right) - \frac{\sqrt{2}Q}{R^2} A_{x,zx} \right] \left(\frac{R^3}{6Q} \right) \quad (2.36)$$

$$C_{yz,yz} = \left[\Theta_{yz} \left(-Q, \frac{\pi}{4}, \frac{\pi}{4} \right) - \Theta_{yz} \left(Q, \frac{\pi}{4}, \frac{\pi}{4} \right) - \frac{Q}{R^2} A_{y,zy} \right] \left(\frac{R^3}{6Q} \right) \quad (2.37)$$

In these equations, $\Theta_{yz} \left(-Q, \frac{\pi}{4}, \frac{\pi}{2} \right)$ for example, represents the yz component of the quadrupole moment when a charge Q is placed R from the origin, with polar angles of $\theta = \frac{\pi}{4}$, $\phi = \frac{\pi}{2}$. For Bishop's results, please refer to Table 8.

Maroulis [Ref14, 21-24] improved Bishop's method. Instead of putting only one point charge, he used several point charges to perturb the system, in order to create a certain electric

field at the origin, where either F_α is zero or $F_{\alpha\beta}$ is zero. As a result, he only needs perturbed energies instead of perturbed multipole moments, in order to calculate various polarizabilities.

3.0 INDUCTION ENERGY ASSOCIATED WITH DPP2 MODEL

The DPP2 model, like the DPP [Ref 2], TTM2 [Ref 20], and AMOEBA [Ref 21] models, adopts the Applequist [Ref 26-28] approach and uses mutually interacting atom-centered point polarizable sites, with Thole-type damping [Ref 15] between the charges and induced dipoles and between the induced dipoles, to describe the polarization interactions. In the DPP, TTM2, and AMOEBA models the values of the atomic polarizabilities were taken from the work of Thole [Ref 15], while the damping coefficients were modified from Thole's work to better fit cluster energies. In the DPP model, the coefficient damping the interactions between the induced dipoles was adjusted so that the model gives three-body energies for the book, prism, cage, and ring isomers of $(\text{H}_2\text{O})_6$ close to those from MP2/aug-cc-pVTZ [Ref 16, 17] calculations.

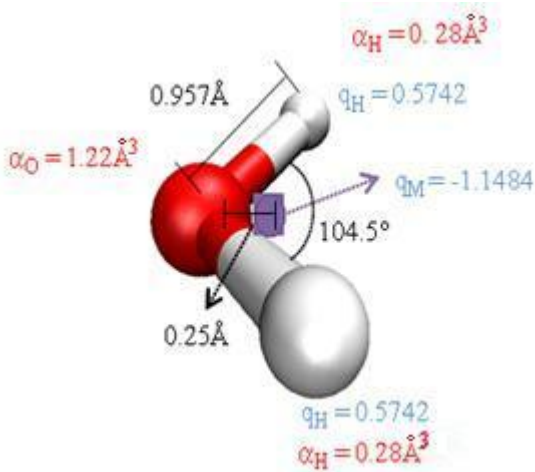


Figure 2 DPP2 water with its parameters

Table 5 Geometry of the H₂O molecule (Angstrom) in the DPP2 model

Atom	x	y	z
O	0.000000	0.000000	-0.065112
H ₁	0.756848	0.000000	0.520901
H ₂	-0.756848	0.000000	0.520901

In the DPP2 model adjusted atomic polarizabilities (keeping the same damping constants as the DPP model) have been adopted to give, simultaneously, the best fit to the atomic polarizability components of the water molecule and the three-body energies (evaluated at the CCSD(T)/aug-cc-pV5Z level [Ref16, 17]) of the four low-lying isomers of the hexamer.

The atomic polarizabilities and damping constants are summarized in Table 6. Interestingly, the resulting values of the atomic polarizabilities are close to those used by Burnham *et al.* in their recently introduced TTM4-F water model [Ref 18]. In calculating the induction energies using the DPP2 model, the electric fields were evaluated using the charges defined in Eq 3.1.

$$q_i^* = 2q_i - \{Z_i - [Z_i - q_i][1 - \exp\left(\frac{\lambda r_{ij}}{\left(\frac{Z_i - q_i}{Z_i}\right)}\right)]\} \quad (2.38)$$

The charge penetration here uses a procedure of Piquemal *et al.* [Ref 29, 30]. However, the inclusion of charge penetration causes only small changes. For example, it leads to only 0.1 kcal/mol in the polarization energies of the water hexamer. So in the remaining part, we do not take charge penetration effects into consideration.

Here we will give a brief description of the procedures used to calculate the polarization energies. The induced dipole μ_i on atom i with polarizability α_i is given by:

$$\mu_i = \alpha_i [\mathbf{E}_i + \sum_{j \neq i} \mathbf{T}_{ij} \mu_j] \quad (2.39)$$

The summation over j involves all other sites within the molecule which including this site i and other molecules which do not contain site i . \mathbf{E}_i is the electric field defined as

$$\mathbf{E}_i = \sum_{k \neq i} f_3(r_{ik}) \frac{q_k \overline{\mathbf{r}_{ik}}}{r_{ik}^2}, \quad (2.40)$$

where the summation over k involves all partial charges on molecules other than the one containing site i , and r_{ik} is the distance between sites i and k .

The dipole tensor \mathbf{T}_{ij} is a 3x3 matrix whose elements are:

$$T_{ij}^{\beta\gamma} = f_5(r_{ij}) \frac{3r_{ij}^{\beta} r_{ij}^{\gamma}}{r_{ij}^5} - f_3(r_{ij}) \frac{\delta_{\beta\gamma}}{r_{ij}^3}, \quad (2.41)$$

where β and γ denote the Cartesian components x , y , or z , and the Thole-type damping functions $f_3(r_{ij})$ and $f_5(r_{ij})$ [Ref 15] are given by

$$f_3(r_{ij}) = 1 - \exp\left[-a \frac{r_{ij}^3}{(\alpha_i \alpha_j)^2}\right] \quad (2.42)$$

and

$$f_5(r_{ij}) = 1 - \left(1 + a \frac{r_{ij}^3}{(\alpha_i \alpha_j)^2}\right) \exp\left[-a \frac{r_{ij}^3}{(\alpha_i \alpha_j)^2}\right] \quad (2.43)$$

Separate values of the damping constant a are employed for the charge-dipole and dipole-dipole interactions. The induced dipoles are solved iteratively, and the induction energies are given by

$$E_{\text{pol}} = -0.5 \sum_i \mathbf{E}_i \mu_i \quad (2.44)$$

Table 6 Parameters in DPP and DPP2 model

Model	Parameter	Value	Unit
DPP	α_{O}	0.837	\AA^3
DPP	α_{H}	0.496	\AA^3
DPP2	α_{O}	1.22	\AA^3
DPP2	α_{H}	0.28	\AA^3
DPP&DPP2	a_{DD} (Dipole-Dipole Damping)	0.3	
DPP&DPP2	a_{CD} (Charge-Dipole Damping)	0.21	
DPP&DPP2	q_{H}	0.5742	e
DPP&DPP2	q_{M}^*	-1.1484	e

*Note: The M site is located on the C_{2v} rotational axis, displaced 0.25 \AA from the O atom towards the H atoms.

4.0 DETAILS OF THE CALCULATIONS

The details of the calculations will be introduced in this chapter. In section 4.1 we introduce how we got the results in Figure 1. In section 4.2 we show the details of our fitting method step by step. In section 4.3, we show the reasons why we cannot get the DPP2 associated A values using this fitting method. In the last section we show how to verify this fitting method.

4.1 HOW TO ESTIMATE THE C CONTRIBUTION:

We use the Orient software developed by Dr. Anthony Stone, etc. [Ref 32] and use the ASP-W4 water force field [Ref 11]. Our system is water dimer and we turn the iteration off. In our calculation first we only use α polarizabilities when calculating the polarization energy E_1 ; then we keep both α and A present and get polarization energy E_2 ; finally we use all α , A and C polarizabilities and get the polarization energy E_3 . So the energy contribution from C term is $E_c = E_3 - E_2$; the energy contribution from A term is $E_A = E_2 - E_1$; the energy contribution from α is E_1 . We performed our calculations at several different O-O distances with results being shown in Figure 1.

4.2 FITING METHOD

There are several approaches to calculate A and C from ab-initio calculations [Ref 5, 6]. For example, Bishop used the perturbed moments with the weak field created by a point charge [Ref 5]; Maroulis used perturbed energies with special charge arrangements [Ref 6]. Based on these two methods we developed a strategy to extract the α , A and C values from DPP2 model. The procedures are as following:

- 1) Put a point charge at different positions and calculate the induction energies

We put a point charge at x, y, z, xy, xz and yz directions separately and at 15Å, 20Å, 25Å, 30Å, 35Å away from the origin. (xy, xz and yz means the bisections of x, y, z axes). The electric field and field gradient at the origin caused by the external point charge is :

$$F_x = \frac{x}{(x^2+y^2+z^2)^{3/2}} \vec{i}, F_y = \frac{y}{(x^2+y^2+z^2)^{3/2}} \vec{j}, F_z = \frac{z}{(x^2+y^2+z^2)^{3/2}} \vec{k} \quad (4.1)$$

$$F_{xy} = \frac{-3xy}{(x^2+y^2+z^2)^{5/2}}, F_{xz} = \frac{-3xz}{(x^2+y^2+z^2)^{5/2}}, F_{yz} = \frac{-3yz}{(x^2+y^2+z^2)^{5/2}},$$

$$F_{xx} = \frac{y^2+z^2-2x^2}{(x^2+y^2+z^2)^{5/2}}, F_{yy} = \frac{x^2+z^2-2y^2}{(x^2+y^2+z^2)^{5/2}}, F_{zz} = \frac{x^2+y^2-2z^2}{(x^2+y^2+z^2)^{5/2}}, \quad (4.2)$$

We can calculate the induction energy of each configuration in step1 using the DPP2 model. The procedures are explained in detail in Chapter 3. We denote the induction energy calculated in each configuration as $E_x, E_y, E_z, E_{xy}, E_{xz}$ and E_{yz} .

2) Do multivariant regression

With equation 2.23 and equations 2.30, 2.31 we obtain:

$$\begin{aligned}
Ez &= E^0 - \frac{1}{2} \alpha_{zz} F_z^2 - \frac{1}{3} A_{z,zz} F_z F_{zz} - \frac{1}{3} A_{z,xx} F_z F_{xx} - \frac{1}{3} A_{z,xx} F_z F_{xx} \\
&\quad - \frac{1}{6} C_{zz,zz} F_{zz} F_{zz} - \frac{1}{6} C_{xx,xx} F_{xx} F_{xx} - \frac{1}{6} C_{yy,yy} F_{yy} F_{yy} \\
&\quad - \frac{1}{3} C_{xx,zz} F_{xx} F_{zz} - \frac{1}{3} C_{yy,zz} F_{yy} F_{zz} - \frac{1}{3} C_{xx,yy} F_{xx} F_{yy}
\end{aligned} \tag{4.3}$$

Since $F_{zz} = -2F_{xx}$, $F_{xx} = F_{yy}$,

$$\begin{aligned}
E_z &= E_{z\alpha} + E_{zA} + E_{zC} \\
&= -\frac{1}{2} F_z^2 \alpha_{zz} \\
&\quad - \frac{1}{3} F_z F_{xx} (A_{z,xx} + A_{z,xx} - 2A_{z,zz}) \\
&\quad - \frac{1}{6} F_{xx} F_{xx} (4C_{zz,zz} + C_{xx,xx} + C_{yy,yy} - 4C_{xx,zz} - 4C_{yy,zz} - 2C_{xx,yy})
\end{aligned} \tag{4.4}$$

Similarly:

$$\begin{aligned}
E_x &= E_{x\alpha} + E_{xA} + E_{xC} \\
&= -\frac{1}{2} F_x^2 \alpha_{xx} - \frac{1}{6} F_{zz} F_{zz} (4C_{xx,xx} - 4C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} + 2C_{yy,zz} + C_{zz,zz})
\end{aligned} \tag{4.5}$$

$$\begin{aligned}
E_y &= E_{y\alpha} + E_{yA} + E_{yC} \\
&= -\frac{1}{2} F_y^2 \alpha_{yy} - \frac{1}{6} F_{zz} F_{zz} (C_{xx,xx} - 4C_{xx,yy} + 2C_{xx,zz} + 4C_{yy,yy} - 4C_{yy,zz} + C_{zz,zz})
\end{aligned} \tag{4.6}$$

$$E_{xy} = E_{xy\alpha} + E_{xyA} + E_{xyC}$$

$$= -\frac{1}{2}F_y^2(\alpha_{xx} + \alpha_{yy}) - \frac{1}{6}F_{xx}F_{xx}(C_{xx,xx} + 2C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} - 4C_{yy,zz} + 4C_{zz,zz} + 36C_{xy,xy}) \quad (4.7)$$

$$\begin{aligned} E_{xz} &= E_{xz\alpha} + E_{xzA} + E_{xzC} \\ &= -\frac{1}{2}F_z^2(\alpha_{xx} + \alpha_{zz}) - \frac{1}{3}F_xF_{xx}(A_{z,xx} + A_{z,zz} - 2A_{z,yy} + 6A_{x,xz}) \\ &\quad - \frac{1}{6}F_{xx}F_{xx}(C_{xx,xx} - 4C_{xx,yy} + 2C_{xx,zz} + 4C_{yy,yy} - 4C_{yy,zz} \\ &\quad \quad \quad + C_{zz,zz} + 36C_{xz,xz}) \end{aligned} \quad (4.8)$$

$$\begin{aligned} E_{yz} &= E_{yz\alpha} + E_{yzA} + E_{yzC} \\ &= -\frac{1}{2}F_z^2(\alpha_{yy} + \alpha_{zz}) - \frac{1}{3}F_yF_{yy}(-2A_{z,xx} + A_{z,zz} + A_{z,yy} + 6A_{y,yz}) \\ &\quad - \frac{1}{6}F_{yy}F_{yy}(4C_{xx,xx} - 4C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} + 2C_{yy,zz} \\ &\quad \quad \quad + C_{zz,zz} + 36C_{yz,yz}) \end{aligned} \quad (4.9)$$

Let

$$Az = A_{z,xx} + A_{z,zz} - 2A_{z,yy}$$

$$Cz = 4C_{zz,zz} + C_{xx,xx} + C_{yy,yy} - 4C_{xx,zz} - 4C_{yy,zz} - 2C_{xx,yy}$$

$$Cx = 4C_{xx,xx} - 4C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} + 2C_{yy,zz} + C_{zz,zz}$$

$$Cy = C_{xx,xx} - 4C_{xx,yy} + 2C_{xx,zz} + 4C_{yy,yy} - 4C_{yy,zz} + C_{zz,zz}$$

$$Cxy = C_{xx,xx} + 2C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} - 4C_{yy,zz} + 4C_{zz,zz}$$

$$Axz = A_{z,xx} + A_{z,zz} - 2A_{z,yy} + 6A_{x,xz}$$

$$Cxz = C_{xx,xx} - 4C_{xx,yy} + 2C_{xx,zz} + 4C_{yy,yy} - 4C_{yy,zz} + C_{zz,zz} + 36C_{xz,xz}$$

$$Ayz = -2A_{z,xx} + A_{z,zz} + A_{z,yy} + 6A_{y,yz}$$

$$C_{yz} = 4C_{xx,xx} - 4C_{xx,yy} - 4C_{xx,zz} + C_{yy,yy} + 2C_{yy,zz} + C_{zz,zz} + 36C_{yz,yz}$$

After we get these induction energies, we perform multivariate regression. In the regression, the Y values are the induction energies for the different configuration. The variants are $-\frac{1}{2}F_{\alpha}^2$, $-\frac{1}{3}F_{\alpha}F_{\alpha\alpha}$, $-\frac{1}{6}F_{\alpha\alpha}F_{\alpha\alpha}$ at different distances as 15Å, 20Å, 25Å, 30Å, 35Å. The regression here helps us to distinguish the induction energies as the sum of energy contributions from α term (E_{α}), A term (E_A) and C term (E_C).

3) Solve a set of linear equations.

After the regression, we get the values of α_{xx} , α_{yy} , α_{zz} directly. We also get the values of C_x , C_y , C_z , C_{xy} , C_{xz} , C_{yz} , A_z , A_{xz} , A_{yz} .

Let us make a matrix M

$$M = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 4 & -4 & -4 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & -4 & 2 & 4 & -4 & 1 & 0 & 0 & 0 \\ 1 & 2 & -4 & 1 & -4 & 4 & 0 & 0 & 0 \\ 1 & 2 & -4 & 1 & -4 & 4 & 36 & 0 & 0 \\ 1 & -4 & 2 & 4 & -4 & 1 & 0 & 6 & 0 \\ 4 & -4 & -4 & 1 & 2 & 1 & 0 & 0 & 36 \end{pmatrix} \quad (4.10)$$

$$\text{Let } C' = (C_{xx:xx} \ C_{xx:yy} \ C_{xx:zz} \ C_{yy:yy} \ C_{yy:zz} \ C_{zz:zz}) \quad (4.11)$$

$$E' = (0 \ 0 \ 0 \ C_x \ C_y \ C_z \ C_{xy} \ C_{xz} \ C_{yz}) \quad (4.12)$$

So $MC = E$. After multiply each side a M^{-1} , we get

$$C = M^{-1}E \quad (4.13)$$

Since M and E are all known, we can now get all the C values.

But we run into problems when we use the same procedures to solve A values. We have five unknown A terms. But we have only four equations. Thus we cannot solve for all A components in this manner. We return to the A issue in the next section.

4.3 Evaluating Dipole-Quadrupole polarizabilities

In section 4.2, it was seen that we need more equations to solve A. Why would this happen? Is it a general result or just unique to C_{2v} symmetry? Can we try to put the point charge in other orientations in order to get enough equations to solve A values? The answer is no. We will see the reasons clearly if we adopt the polar expressions.

The dipole expression for the field and field gradient are:

$$F_x = \frac{\sin \theta \cos \phi}{R^2}, F_y = \frac{\sin \theta \sin \phi}{R^2}, F_z = \frac{\cos \theta}{R^2}, \quad (4.14)$$

$$F_{xx} = \frac{\sin^2 \theta \sin^2 \phi + \cos^2 \theta - 2 \sin^2 \theta \cos^2 \phi}{R^3},$$

$$F_{yy} = \frac{\sin^2 \theta \cos^2 \phi + \cos^2 \theta - 2 \sin^2 \theta \sin^2 \phi}{R^3},$$

$$F_{zz} = \frac{\sin^2 \theta - 2 \cos^2 \theta}{R^3},$$

$$F_{xz} = \frac{-3 \sin \theta \cos \phi \cos \theta}{R^3},$$

$$F_{yz} = \frac{-3 \sin \theta \sin \phi \cos \theta}{R^3}, \quad (4.15)$$

Let us express induction energy from A term as E_A . We have

$$\begin{aligned}
 -3R^5 E_A = & \\
 2A_{x,zx} * (-3 \sin \theta \cos \phi \sin \theta \cos \phi \cos \theta) + 2A_{y,zy} * (3 \sin \theta \sin \phi \sin \theta \sin \phi \cos \theta) + A_{z,xx} * & \\
 \cos \theta (\sin^2 \theta \sin^2 \phi + \cos^2 \theta - 2 \sin^2 \theta \cos^2 \phi) + A_{z,yy} * \cos \theta (\sin^2 \theta \cos^2 \phi + \cos^2 \theta - & \\
 2 \sin^2 \theta \sin^2 \phi) + A_{z,zz} * \cos \theta (\sin^2 \theta - 2 \cos^2 \theta) & \quad (4.16)
 \end{aligned}$$

Let

$$C_1 = -3 \sin^2 \theta \cos \theta \cos^2 \phi,$$

$$C_2 = -3 \sin^2 \theta \cos \theta \sin^2 \phi,$$

$$C_3 = \cos \theta,$$

$$C_4 = \cos \theta (1 - 3 \cos^2 \theta),$$

Then equation 4.15 becomes

$$\begin{aligned}
 -3R^5 E_A = 2C_1 A_{x,zx} + 2C_2 A_{y,zy} + (C_1 + C_3) A_{z,xx} + (C_2 + C_3) A_{z,yy} + C_4 A_{z,zz} & \\
 = C_1 (2A_{x,zx} + A_{z,xx}) + C_2 (2A_{y,zy} + A_{z,yy}) + C_3 (A_{z,xx} + A_{z,yy}) & \\
 + C_4 A_{z,zz} & \quad (4.17)
 \end{aligned}$$

Let

$$2A_{x,zx} + A_{z,xx} = E_{A1} \quad (4.18)$$

$$2A_{y,zy} + A_{z,yy} = E_{A2} \quad (4.19)$$

$$A_{z,xx} + A_{z,yy} = E_{A3} \quad (4.20)$$

$$A_{z,zz} = E_{A4} \quad (4.21)$$

We know that $A_{z,xx} + A_{z,yy} + A_{z,zz} = 0$, so after we know equation 4.20, equation 4.21 adds no more valuable information here. For the three equations 4.18~4.20, we have four

unknowns. As a result, we can only get the value of $A_{z,zz}$ and the sum of other A components. We cannot get the individual $A_{x:zx}$, $A_{y:zy}$, $A_{z:xx}$, $A_{z:yy}$ values through this fitting method.

4.4 HOW TO VERIFY THIS METHOD.

So how good is this fitting method? We used ASP-W4 water force field to get all the induction energies we need in different distances and different directions. Then we calculate all the C values out using the procedures described in section 4.2 and compare them with their original values in ASP-W4 force field (polarizabilities in ASP-W4 are written in spherical expression and origin at Oxygen. We convert them into Cartesian expression here). As we can see from Table 7, the results we get from this fitting method are almost the same as their original values. This assures us that our fitting method works well.

Table 7 C values from ASP –W4 and our fitting method^a

	$C_{xx,xx}$	$C_{xx,yy}$	$C_{xx,zz}$	$C_{yy,yy}$	$C_{yy,zz}$	$C_{zz,zz}$	$C_{xy,xy}$	$C_{xz,xz}$	$C_{yz,yz}$
ASP-W4	11.88	-6.81	-5.07	11.73	-4.92	9.99	7.79	13.02	7.02
FM	11.93	-6.83	-5.10	12.10	-4.86	9.96	7.86	13.14	7.04

a: Units are in atomic units; origin is at center of mass.

5.0 RESULTS AND DISCUSSION

The associated multipole polarizabilities of DPP and DPP2 differ appreciably from the ab-initio values given by Bishop. We tried different ways in order to improve the results. Changing the geometry does not help much. When we change the polarizable site from oxygen to M site, results are improved to some extent. But the signs of three C components are still incorrect. The most exciting trial is when we split the polarizability of oxygen onto its two electron lone pairs. The results are much improved. We give our analysis on why such changes on the location of polarizability make the values in better agreement with ab-initio values based on a single center expansion.

5.1 DPP AND DPP2

As we can see from Table 8, DPP2 performs better than DPP. But neither is high successful at accounting for C. Not only is $C_{xx:zz}$ in DPP more than five times bigger than the ab-initio result, but also three components in both the DPP and DPP2 models have the wrong sign.

Table 8 Values of C components of DPP1 and DPP2^a

	$C_{xx,xx}$	$C_{xx,yy}$	$C_{xx,zz}$	$C_{yy,yy}$	$C_{yy,zz}$	$C_{zz,zz}$	$C_{xy,xy}$	$C_{xz,xz}$	$C_{yz,yz}$
DPP	38.39	-9.87	-28.53	-7.49	17.36	11.17	9.78	40.82	-9.28
DPP2	19.45	-5.25	-14.20	-2.39	7.63	6.57	4.52	22.67	-3.74
DPP2G	19.49	-5.25	-14.24	-2.40	7.65	6.59	4.53	22.74	-3.75
Bishop[3]	11.78	-6.85	-4.93	12.02	-5.17	10.10	8.48	11.06	4.72

a: Units are in atomic units. Origin is at center of mass.

5.2 CHANGE THE GEOMETRY

As we can see in Tables 4 and 5, there are some differences between water geometry used in DPP2 and that used by Bishop. Could such differences lead to significant differences in the C values? Another calculation using the water geometry used by Bishop shows that this is not the case. The results are named DPP2G in Table 8.

5.3 CHANGE POLARIZABILITY FROM OXYGEN TO M SITE

When using distributed charges to describe the multipole moments of water molecule, it is found that though we can fit the individual charges to get very good molecular dipole moments which are closed to experimental values, but that is not possible to get the

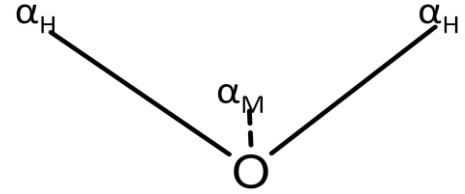


Figure 3 Schematic structure of DPP2M model

quadrupole moments right at the same time. After the charge is moved from oxygen to the M site, this problem is solved. So we try the same strategy here to try to achieve good α and C values at the same time. In DPP2M, we changed the polarizability from oxygen to M site. In DPP2M2, we changed the polarizability on M site from 1.22 to 1.36 Å³ and the polarizability on hydrogen from 0.28 to 0.22 Å³. The results are improved by some extent, but not much. Some values are even worse. The sign errors still exist.

Table 9 α and C values of polarizable M site DPP2 model^a

	Bishop	DPP2	DPP2M	DPP2M2
α_{xx}	9.24	10.51	10.20	10.48
α_{yy}	7.91	9.20	9.07	9.59
α_{zz}	8.55	9.70	9.21	9.71
$C_{xx,xx}$	11.78	19.45	18.29	13.24
$C_{xx,yy}$	-6.85	-5.25	-3.78	-2.25
$C_{xx,zz}$	-4.93	-14.20	-14.51	-10.99
$C_{yy,yy}$	12.02	-2.39	-2.82	-1.56
$C_{yy,zz}$	-5.17	7.63	6.60	3.81
$C_{zz,zz}$	10.10	6.57	7.92	7.18
$C_{xy,xy}$	8.48	4.52	3.54	2.05
$C_{xz,xz}$	11.06	22.68	21.54	16.76
$C_{yz,yz}$	4.72	-3.74	-2.90	-1.15

a: Units are in atomic units; origin is at the center of mass

5.4 FOUR-SITE POLARIZABLE MODELS

When we were using ASP-W4 model to verify our fitting method, we noticed that the E_{yC} in equation 4.6 is positive for ASP-W4. However in DPP2 and in the other models above, E_{yC} are negative. The sign differences in E_{yC} lead to the sign errors in $C_{yy:yy}$, $C_{yy:zz}$ and $C_{yz:yz}$. So how can we improve our model in order to make E_{yC} positive?

We know that the electron lone pairs on the oxygen atom point out of the water plane. This suggests making a model with out-of-plane dipole polarizable sites.

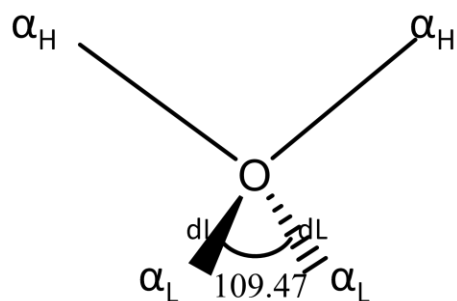


Figure 4 Schematic structure of 4-site polarizable model

In the CPE2 model [Ref 19], the charge distribution is represented by a Gaussian basis function on each atom plus two dipole-like basis functions on the sites of the oxygen lone pairs.

The lone-pair charge distributions are placed 0.65\AA away from the oxygen

nucleus, forming a dipole-oxygen-dipole angle of 109.47° . They apply this strategy to charge distributions. But can we also move polarizabilities to the electron lone pairs? In model “lone pair 4 site polarizable model 1” (LP4P1), we split the polarizability 1.22\AA^3 of oxygen to 0.66\AA^3 on each lone pair. The positions of the lone-pair polarizabilities are the same as the positions of

lone-pair charge distributions in CPE2 model. The distance from lone pair site to the oxygen is $d_L=0.61\text{\AA}$. The α_{L1} -Oxygen- α_{L2} angle is 109.47° . As seen from in Table 11, all nine C components of LP4P1 now have correct signs as compared to the ab-initio results. The values are also much closer to Bishop's. We also tried other parameters of such four polarizable sites model in order to compare the agreement with Bishop's results. These sets of parameters used in models named from LP4P2 to LP4P7 are listed in Table 10. The results are listed in Table 11. Among these seven sets of parameters, LP4P7 give us the best results overall.

Table 10 Parameters used in seven lone-pair 4-site polarizable models^a.

Model	LP4P1	LP4P2	LP4P3	LP4P4	LP4P5	LP4P6	LP4P7
α_H	0.28	0.28	0.28	0.28	0.22	0.22	0.22
α_L	0.61	0.5	0.5	0.61	0.68	0.61	0.61
d_L	0.61	0.61	0.5	0.5	0.61	0.61	0.5

a: Units are in atomic units; origin is at the center of mass.

Table 11 α and C values of various 4 site polarizable DPP2 model

	Bishop	DPP2	LP4P1	LP4P2	LP4P3	LP4P4	LP4P5	LP4P6	LP4P7
α_{xx}	9.24	10.51	10.18	9.23	9.03	9.94	10.14	9.49	9.30
α_{yy}	7.91	9.20	9.47	8.51	7.79	8.70	9.70	9.06	8.32
α_{zz}	8.55	9.70	10.32	9.30	8.84	9.81	10.48	9.80	9.35
$C_{xx,xx}$	11.78	19.45	16.56	17.75	19.16	18.19	10.81	11.54	13.25
$C_{xx,yy}$	-6.85	-5.25	-9.61	-8.92	-8.00	-8.39	-8.72	-8.26	-7.12
$C_{xx,zz}$	-4.93	-14.20	-6.95	-8.84	-11.17	-9.80	-2.09	-3.28	-6.13
$C_{yy,yy}$	12.02	-2.39	16.44	12.01	7.18	10.43	20.30	17.49	11.43
$C_{yy,zz}$	-5.17	7.63	-6.83	-3.10	0.82	-2.04	-11.58	-9.23	-4.31
$C_{zz,zz}$	10.10	6.57	13.78	11.93	10.34	11.84	13.66	12.51	10.44
$C_{xy,xy}$	8.48	4.52	7.98	7.14	6.23	6.74	7.66	7.08	5.76
$C_{xz,xz}$	11.06	22.68	23.96	24.35	25.17	24.88	18.32	18.56	19.63
$C_{yz,yz}$	4.72	-3.74	8.86	5.92	2.35	4.61	11.92	10.10	5.79

Note: Units are in atomic units; origin is at center of mass.

6.0 CONCLUSION

In conclusion, we show that our fitting method can be treated as a general method for mapping distributed polarizabilities into multipole polarizabilities based on single center expansion. It is very accuracy and simple in reproducing the C components. Here we used the multivariant regression instead of using several external point charges to distinguish the energy contributions from α , A and C terms. From the better agreement of our 4-site polarizable model with ab-initio results in C components' values, we also show that adding polarizabilities on electron lone pairs of oxygen is necessary for developing more accurate water force field using distributed polarizabilities. In future work, one should test our four-site polarizable model on the water dimer and on other water clusters to see whether it can describe accurately the induction energies of these systems.

REFERENCES

- 1, R. Kumar, F. F. Wang, G. Jenness, and K. D. Jordan, *J. Chem. Phys.*, **132**, 014309(2010).
- 2, A. Defusco, D. P. Schofield, and K. D. Jordan, *Mol. Phys.* **105**, 2681(2007)
- 3, D. M. Bishop, J. Pipin, *Theor. Chim. Acta.*, **71**, 247(1987)
- 4, W. Murphy, *J. Chem. Phys.*, **67**, 5877(1977)
- 5, C. Millot, J.-C. Soetens, M.T.C. Martins Costa, M.P. Hodges, A. J. Stone, *J. Phys. Chem. A* **102**, 754(1998)
- 6, “The Theory of Intermolecular Forces”, A. J. Stone, Oxford University Press, 1996
- 7, K. Kitaura and K. Morokuma, *Int. J. Quantum Chem.* **10**, 325(1976)
- 8, B. Jeziorski, R. Moszynski and K. Szalewics, *Chem. Rev.*, **94**, 1887(1994)
- 9, P. S. Bagus and F. Illas, *J. Chem.Phys.* **96**, 8962(1992)
- 10, W. J. Stevens and W. H. Fink, *Chem. Phys. Lett.* **139**, 15(1987)
- 11, M. P. Hodges and A. J. Stone, *J. Phys. Chem. A*, **101 (48)**, 9163(1997)
- 12, A.D. McLean, M. Yoshimine, *J. Chem. Phys.*, **47**, 1927(1967)
- 13, Buckingham A.D. (1967) In: Hirschfelder J.O. (ed) *Advanced in Chemical Physics*, Vol 12. Interscience, New York, p 107
- 14, G Maroulis, A. J. Thakkar, *J. Chem. Phys.*, **88**, 7623(1988)
- 15, B.T. Thole, *J. Chem. Phys.* **59**, 341(1981)
- 16, J. Thom, H. Dunning, *J. Chem. Phys.* **90**, 1007(1989)

- 17, R. A. Kendall, J. Thom, H. Dunning, and R. J. Harrison, *J. Chem. Phys.*, **96**, 6796(1992)
- 18, C. J. Burnham, D. J. Anick, P. K. Mankoo, and G. F.Reiter, *J. Chem. Phys* **128**, 154519(2008)
- 19, J.-P. Piquemal, R. Chelli, P. Procacci and N. Gresh, *J. Phys. Chem. A*, **111**, 8170(2007)
- 20, C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.*, **116**, 1500(2002)
- 21, P. Ren and J. W. Ponder, *J. Phys. Chem. B*, **107**, 5933(2003)
- 22, G. Maroulis, *J. Phys. B: At. Mol. Opt. Phys*, **26**, 2957(1993)
- 23, G. Maroulis, *J. Chem. Phys*, **108**, 5432(1998)
- 24, G. Maroulis, A.J. Thakkar, *J. Chem. Phys.*, **93**, 4164(1990)
- 25, E. R. Batista, S. S. Xantheas, H. Jónsson, *J. Chem. Phys.*, **109**, 4546(1998)
- 26, J. Applequist, *J.Math.Phys*, **24**, 736(1983)
- 27, J. Applequist. *J.Chem.Phys.* **83**, 809(1995)
- 28, J. Applequist , Carl, J. R. and Fung, K. -K., *J. Amer. Chem. Soc.* **94**, 2952(1972)
- 29, J. -P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gersh, and T. A. Darden, *J. Chem. Phys.* **124**, 104101(2006)
- 30, G. A. Cisneros, S. N. -I. Tholander, O. Rarisel, T. A. Darden, D. Elking, L. Perera, and J. -P. Piquemal, *Int. J. Quant. Chem.* **108**, 1905(2008)
- 31, T. Pluta, J. Noga, and R. J. Bartlett, *Inter. J. Quan. Chem.* **28**, 379(1994)
- 32, A. J. Stone, A. Dullweber, O. Engkvist, E. Frascini, M. P. Hodges, A. W. Meredith; D. R. Nutt, P. L. A. Popelier and D. J. Wales (2002) 'Orient: a program for studying interactions between molecules, version 4.5,' University of Cambridge, Enquiries to A. J. Stone, ajs1@cam.ac.uk.