

MOLECULAR DYNAMICS STUDIES ON NEURAL
NETWORK *AB INITIO* POTENTIAL ENERGY
SURFACES

By

HUNG M. LE

Bachelor of Science in Chemistry

University of Central Oklahoma

Edmond, Oklahoma

2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2009

MOLECULAR DYNAMICS STUDIES ON NEURAL
NETWORK *AB INITIO* POTENTIAL ENERGY
SURFACES

Thesis Approved:

Dr. Lionel M. Raff

Thesis Adviser

Dr. Mark G. Rockley

Dr. Jeffery L. White

Dr. Ranga Komanduri

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to thank specially my research advisor, Dr. Lionel M. Raff, for being helpful during the course of my research at Oklahoma State University. It is a special honor to be his last graduate student.

I also thank Dr. Mark. G. Rockley, one of my graduate committee member and my nice neighbor, for giving me some helpful comments. Thanks to Dr. Jeffery L. White and Dr. Ranga Komanduri for their willingness to serve in the committee.

Half of this work is dedicated to my mom. The other half of this work is dedicated to a very special person, whose name is forever in my heart, my plan, and my future.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
II. REVIEW OF POTENTIAL FITTING METHODS	10
Shepard interpolation method.....	10
Interpolating moving least squares (IMLS) method.....	13
Neural network (NN) method.....	15
Support vector machine (SVM) method.....	20
III. <i>CIS-TRANS</i> ISOMERIZATIONS AND N-O BOND DISSOCIATION OF NITROUS ACID (HONO).....	23
Development of <i>ab initio</i> potential surface.....	23
Electronic structure calculations.....	23
Modified novelty sampling procedure.....	27
Normal mode analysis on the new PES	34
Neural network committee.....	35
Trajectory calculations.....	37
<i>Cis-trans</i> isomerizations	39
N-O bond dissociation	44
Summary.....	48
IV. INVESTIGATION OF $BeH + H_2 \rightarrow BeH_2 + H$	51
Development of <i>ab initio</i> potential surface.....	51
Novelty sampling of configurations.....	51
NN fitting of BeH_3 potential.....	57
Classical dynamics investigation of $BeH + H_2 \rightarrow BeH_2 + H$	61
Initializing the $BeH + H_2$ reaction	61
Classical dynamics investigation	65
Summary.....	71

Chapter	Page
V. MOLECULAR DISSOCIATION OF HYDROGEN PEROXIDE (HOOH)	74
Introduction of a new sampling method	74
Electronic structure calculations	83
Fitting methods	83
Molecular dissociation of HOOH: A proof of internal hydrogen bonding.....	90
Summary	99
VI. CONCLUSION.....	102
REFERENCES	107

LIST OF TABLES

Table	Page
III.1 Equilibrium energies and potential barriers of HONO	27
III.2 Number of add configurations during HONO sampling.....	33
III.3 Equilibrium, maximum, and minimum atomic distances of HONO	33
III.4 Experimental and theoretical wave numbers of HONO	35
III.5 Average absolute errors of NN potential surfaces of HONO	37
III.6 <i>Cis-trans</i> isomerization rate constants	43
III.7 N-O bond dissociation rate constants	47
IV.1 Number of add configurations during BeH ₃ sampling	57
IV.2 Average absolute errors of NN potential surfaces of BeH ₃	60
IV.3 Fitting parameters for BeH potential surface.....	62
IV.4 Fitting parameters for H ₂ potential surface.....	64
IV.5 Reaction probability and cross section of $BeH + H_2 \rightarrow BeH_2 + H$	70
V.1 Number of configurations obtained from partial optimization	78
V.2 Theoretical wave numbers of HOOH from various calculations.....	81
V.3 Number of configurations obtained from HOOH sampling process	82
V.4 Average absolute errors of NN potential surfaces of HOOH	84
V.5 First-order rate constants at different levels of internal energy	94
V.6 First-order rate constants at 3.4 eV obtained on NN surfaces trained with different sets of data.....	95

LIST OF FIGURES

Figure	Page
I.1 <i>Cis</i> and <i>trans</i> configurations of HONO.....	5
I.2 BeH ₃ system with six atomic distance assigned from r ₁ →r ₆	6
I.3 Equilibrium configuration of HOOH	9
II.1 Structure of a two-layer feed-forward NN with 34 neurons	17
II.2 An illustrative example of over-fitting.....	18
II.2 Mean squared error of fitting, validation, and testing sets	18
II.3 Plots of output vs. targets for training, validation, and testing sets.....	19
II.4 Example of an SVM regression using radial basis kernel.....	21
III.1 Rotational barrier of HONO	25
III.2 N-O dissociation barrier of HONO.....	26
III.3 Non-normalized probability distribution of configurations in the second iteration	31
III.4 Non-normalized probability distribution of configurations in the last iteration.....	32
III.5 First-order decay plot of <i>cis</i> → <i>trans</i> isomerization when exciting OH mode at 1.7 eV internal energy.....	41
III.6 First-order decay plot of <i>trans</i> → <i>cis</i> isomerization when exciting OH mode at 1.7 eV internal energy.....	42
III.7 First-order decay plot of N-O bond dissociation when exciting OH mode at 3.1 eV internal energy.....	46

Figure	Page
IV.1 Distribution of minimum distances of the first 1,300 BeH ₃ points	54
IV.2 Distribution of average distances of the first 1,300 BeH ₃ points	55
IV.3 Absolute error distribution of all data of NN committee (BeH ₃)	59
IV.4 Distribution of gradient testing absolute error.....	61
IV.5 Potential energy surface of BeH	63
IV.6 Potential energy surface of H ₂	65
IV.7 Reaction probability vs. various translational energies of <i>BeH + H₂ → BeH₂ + H</i>	67
IV.8 Plot of reaction cross section vs. various translational energies for <i>BeH + H₂ → BeH₂ + H</i>	69
V.1 Plot of targets vs. training outputs for all HOOH data points.....	85
V.2 Distribution of the absolute testing error for NN committee	86
V.3 Various SVM training with ρ=0.03	88
V.4 Potential barrier of O-O dissociation	91
V.5 First-order decay plot of O-O dissociation at 3.4 eV.....	93
V.6 Rice-Ramsperger-Kassel plot of HOOH dissociation	96
V.7 Snapshots of HOOH dissociation	98

CHAPTER I

INTRODUCTION

Born-Oppenheimer potential-energy surfaces (PES) have been a subject of study for decades. Since it was first proposed, many molecular dynamics (MD) studies have been conducted based on the most traditional technique using empirical and semi-empirical PES's. These surfaces have very limited accuracy in terms of both potential energy fitting and chemical dynamics descriptions. The idea of fitting energies resulting from electronic structure calculations (*ab initio* energies) has been in use since the early 80s with many limitations mainly because of huge computational cost.

Morse potential function¹ is a widely-employed method to describe chemical bonds. Potential energy functions of this type are obtained from fitting experimental data, such as chemical bond enthalpies, molecular vibrational energies, and equilibrium bond distances. Sudhakaran and Raff² made an effort in 1984 to describe a three-atom system for the reaction $H(D) + HBr(DBr)$ by using a combination of three Morse potentials. Systems with triatomic molecules and higher complexity usually require bending potentials to describe the energy between chemical bonds and torsional potentials to describe rotations about chemical bonds. Guan and Thompson reported two different PES's for *cis*-HONO and *trans*-HONO,³ and studied the *cis-trans* isomerization as well as the dissociation of the N-O bond. In a latter study, Guo and Thompson⁴ generalized

the *cis* and *trans* potential surfaces into a unique form that described both configurations by a method named “valence bond potential.”

Fitting *ab initio* potential energies of a molecule using an empirical or semi-empirical surface requires considerable mathematical and computational effort. Back in the 1980s, this approach was very time consuming and inapplicable. Water is one of the most important molecules that supports living organisms. Consequently, it has been studied extensively. The reaction $H_2O \rightarrow OH + H$ was investigated theoretically by Wright and Shih⁵ when they constructed an *ab initio* PES based on self-consistent-field⁶ (SCF) and multi-reference double excitation configuration-interaction^{7,8} (CI) calculations on the double-zeta basis set (cc-pVDZ).⁹⁻¹³ Vinyl bromide (CH_2CHBr) is a very complex six-atom molecule for MD studies as it has 12 internal coordinates and six open reaction channels. A semi-empirical surface that fits the *ab initio* potential energies of this molecule was first reported by Rahaman and Raff.¹⁴ Another example of semi-empirical PES's as fitting functions of *ab initio* potential energy was presented by Kuhn *et al.*¹⁵ In their work, four different PES fits were reported for hydrogen peroxide (HOOH), to *ab initio* calculations which were performed by employing density functional theory (DFT) using the hybrid Beck 3 parameter Lee-Yang-Parr (B3LYP) calculations,^{16,17} and second-order complete active space calculations (CASPT2)^{18,19}, a very expensive *ab initio* method, on the triple-zeta basis set (cc-pVTZ).⁹⁻¹³ Four different PES fits were reported in that study, and each fit required more than 6,000 *ab initio* points.

Currently, MD studies have been extended further since the development of some effective fitting techniques. Several molecules and chemical reactions with different levels of complexity have been investigated using these new techniques. The Shepard

interpolation method which was first proposed by Ischtwan and Collins attempts to fit *ab initio* data automatically.²⁰ In this method, the potential energy at a given configuration is expressed as a summation of weighted Taylor expansions around N chosen points in the database. During MD trajectories, the fitting process has to be re-analyzed to determine the weights of each Taylor expansion. Therefore, this method does not provide an analytical function. Interpolating moving least-squares (IMLS) method is another technique that was first introduced by Maisuradze *et al.*²¹ The fitted potential in this method is a linear combination of many basis functions that are constructed by least-squares fitting of data points in hyperspace. This method has been shown to be practical in MD studies, and in fact, it has been employed to investigate several chemical reactions theoretically. Neural networks (NN) are a robust and powerful fitting method, which have been developed for many decades.²² In this method, after a sufficient number of data points is obtained (the database required when using a NN is usually much larger than the database requirements for other methods), an analytical function is employed to fit them only once. This brings considerable computational advantages as retrieving the database during MD trajectories is not required. In our investigations, three different studies have been made with various approaches using this fitting technique. The last fitting technique that is also employed during the course of this research is the regression fitting built into the support vector machine (SVM) method, a “cousin” tool of NN, which was first proposed by Cortes and Vapnik.²³ In the area of MD studies, this technique has not been widely used. More details about each individual technique will be discussed in detail in a latter section.

The first study that we have conducted using NN fitting is the *cis-trans* isomerizations and N-O bond rupture of nitrous acid (HONO).²⁴ HONO has two equilibrium minima in term of potential energy. The *cis* configuration has a dihedral angle of 0°, while the *trans* configuration has a dihedral angle of 180° (HONO is perfectly planar in both equilibrium configurations).

The unimolecular reaction dynamics of HONO has been investigated since 1982 both theoretically and experimentally. In an experimental work, McDonald and Shirk²⁵ studied the *cis-trans* isomerizations in solid N₂ and Ar matrices, and the rate of the reaction was not first order according to the reported results. In a latter study conducted by Shirk and Shirk,²⁶ the rate of *cis*→*trans* reaction was reported to be almost 100 times faster than the reverse reaction. A similar reaction was investigated in various matrices (H₂S, NO₂, and Kr) by Khriachtchev *et al.*²⁷ The first reported theoretical study of the gas phase reaction of HONO was conducted by Guan, Lynch, and Thompson²⁸ using an empirical surface, and the reactions (*cis-trans* isomerization and N-O dissociation) were found to be first order. In a latter work, Guan and Thompson³ extended the theoretical investigations, and reported two PES's for *cis* and *trans* HONO. They concluded that the photo-excitation effect was mode-specific as different excitations on six vibrational modes were executed for MD investigations. Also, *cis*→*trans* isomerization was found to be much faster than the reverse reaction. However, the characteristic of these two individual PES's do not generalize the dissociation process well because the molecule tends to switch very fast from *cis* to *trans* and vice versa. Thus, in 2003, Guo and Thompson⁴ developed a PES that describes both configurations, which was known as the “valence bond potential surface.” In our study of HONO, we attempted to collect data

points that describe both *cis* and *trans* configurations as well as N-O bond dissociation, and employed a unique NN to fit the data and produce one PES. Theoretical investigations of the two reactions (isomerization and dissociation) were executed with different vibrational excitations for some specific internal energy levels.



Figure I.1 *Cis* and *trans* configurations of HONO

Developing *ab initio* NN PES's that describe complex molecular dissociations has been proven to be practical. In a theoretical investigation of vinyl bromide (CH₂CHBr) dynamics, Doughan *et al.*²⁹ and Malshe *et al.*³⁰ constructed a PES that describes six different reaction channels when vinyl bromide is vibrationally excited. The configurations in the vinyl bromide database are obtained using a “novelty sampling” (NS) technique first introduced by Raff *et al.*,³¹ which is based upon MD trajectories. From the studies of HONO and CH₂CHBr, the capability of constructing PESs to describe dissociations of those molecules with 6 atoms or less is established. Our second effort to improve the NN method is to develop a NN PES that describes a collision in very wide hyperspace between two diatomic molecules. In our second study, the formation of beryllium hydride (BeH₂) due to the collision of BeH and H₂ is investigated.³² The six-dimensional configuration hyperspace of this four-body system is believed to be more complicated than the hyperspace of HONO, since the extension of

atomic distances is considerably larger. To obtain data points, we employed similar sampling procedure as previously used for HONO to sample BeH_3 configurations. Once the PES is constructed, we employ classical trajectories to calculate the reaction probability and cross sections.

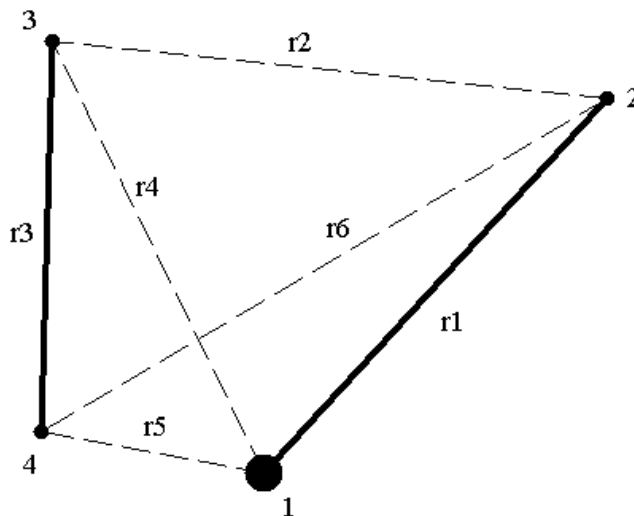


Figure I.2 BeH_3 system with six atomic distances assigned as $r_1 \rightarrow r_6$

In an earlier investigation, Collins and Zhang studied the reaction of $\text{BeH} + \text{H}_2 \rightarrow \text{BeH}_2 + \text{H}$ using the Shepard interpolation method to fit 1,300 data points.³³ Molecular dynamics was executed using a quantum reactive scattering method developed by Zhang and co-workers,³⁴⁻³⁶ and the maximum reaction probability at zero total angular momentum was reported as 0.11. Prior to this work, they performed another PES fit for BeH_3 system using the same technique with only 438 configurations,³⁷ and the reaction was investigated using quasi-classical dynamics. Koput and Peterson³⁸ reported an *ab initio* prediction of BeH_2 (and the similar compounds with deuterium substituted

for hydrogen) potential surface and vibration-rotation energy levels with high precision using multiple electronic structure calculations (MP2, MP4(SDQ),³⁹⁻⁴⁴ CCSD⁴⁵⁻⁴⁸). Our study of BeH₃ is considered as an extended work of a previous study³³ because the set of 1,300 data points provided by Collins and Zhang³³ is used to construct a temporary potential surface as part of the modified novelty sampling procedure. Thereafter, similar electronic structure calculations (MP2 level of accuracy applied on the 6-311G(d,p) basis set) are executed to calculate the potential energy of the generated configurations. After constructing the NN PES, the collision is investigated at various translational energy levels (from 0.415 eV to 0.829 eV) using classical dynamics.

The effectiveness of the “novelty sampling” procedure³¹ has not yet fully investigated. This procedure is based mainly on MD trajectories that sample data points in hyperspace. To evaluate the necessity of employing MD trajectories during configuration sampling, a new sampling method named “gradient sampling” is introduced. Briefly, a “gradient sampling” method is employed to search for regions that requires more configurations by analyzing the NN-fitting gradients (with respect to the inputs). The success of this method would prove MD sampling is not obligatory in any potential surface developments. We have chosen hydrogen peroxide (HOOH), a widely-studied four-body molecule, as an illustrative example for the third NN PES study. The aim of the last study is to develop a surface that fully describes O-O bond rupture.

Hydrogen peroxide is an important gas in atmospheric chemistry with a strong oxidation property. It has been researched in many theoretical and experimental works. Rizzo, Hayden, and Crim used direct excitation of overtone vibrations to study the reaction dynamics of HOOH.⁴⁹ This technique was also used by Dubal and Crim when

overtone vibration excitation was initiated as a function of the wavelength.⁵⁰ Laser induced fluorescence detection was used by Ticich *et al.*⁵¹ to detect the fragments of OH radicals during the dissociation process caused by exciting the OH stretching mode. Several other experimental investigations were conducted as the dissociation was monitored during the excitation of one or more vibrational quantum modes.⁵²⁻⁵⁸ Lin and Guo⁵⁹ executed full-dimensional quantum calculations using the Lanczos algorithm on an *ab initio* PES with high quality electronic structure calculations. The work was further extended when they investigated HOOH and its deuterated isotopomers.⁶⁰ As introduced above, four semi-empirical surfaces of HOOH were developed by Kuhn *et al.*¹⁵ with more than 6,000 data points. This is a remarkable effort of using semi-empirical functions to fit so many data points. Guo *et al.*⁶¹ and Maisuradze *et al.*⁶² performed a fit for the PCPSDE surface reported by Kuhn *et al.*¹⁵, and the O-O bond rupture was investigated on both the PCPSDE and IMLS fitted surfaces. The reported results show a good agreement as they fit the first-order rate coefficients to the Rice-Ramsperger-Kassel equation, and the *s* values are in good agreement (3.4 from MD investigations on PCPSDE surface and 3.2 from MD investigations on the IMLS surface).

Unlike nitrous acid (HONO), the equilibrium configuration of hydrogen peroxide (HOOH) is neither *cis* nor *trans* because each oxygen atom has two lone pairs of electrons. According to MP2 optimization on the 6-31G* basis set, at equilibrium, the dihedral angle is approximately 121°, while the HOO angle is 98.6°. The HO and OO bond distances are 0.98 Å and 1.47Å, respectively.

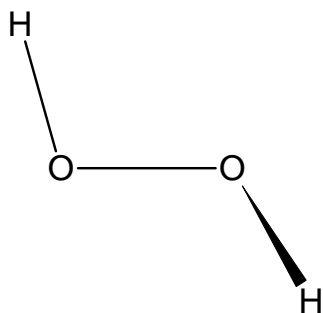


Figure I.3 Equilibrium configuration of HOOH.

An important characteristic of HOOH is the ability to form internal hydrogen bonding. As we shall see later in our study,⁶³ during the dissociation process, it will be shown that one or both hydrogen atoms have a tendency to move closely to the other oxygen, and this formation of hydrogen bond resists O-O rupture strongly. This property was previously proposed in a theoretical study by Harding.⁶⁴

CHAPTER II

REVIEW OF POTENTIAL FITTING METHODS

Sampling and fitting data are two important processes that are executed simultaneously during the development of an *ab initio* PES. Since the sampling procedure varies from study to study to enhance the efficiency due to the nature of the molecular system, in this chapter, we assume that the database is already in-hand, and introduce some well-known fitting methods.

I. Shepard interpolation method

This method was first proposed by Ischtwan and Collins.²⁰ Basically, the potential energy at a given configuration during MD trajectories is a sum of many weighted Taylor expansions around the selected configurations stored in the database. The form of PES's, as this method is applied, is not an analytical mathematical function.

For a four-body system, let us denote vector $R = \{R_1, R_2, R_3, R_4, R_5, R_6\}$ as the atomic distance vector. In practice, the input vector is the inverse atomic distance vector, denoted as Z ($Z_i = 1/R_i$). The Taylor expansion up to the second order terms at a configuration in the database is expressed as

$$T_i(Z) = V(Z(i)) + \sum_{j=1}^6 [Z_j - Z_j(i)] \frac{\partial V}{\partial Z_j} + \frac{1}{2} \sum_{j=1}^6 \sum_{k=1}^6 [Z_j - Z_j(i)] \times [Z_k - Z_k(i)] \frac{\partial^2}{\partial Z_j \partial Z_k}. \quad (1)$$

The potential at a point that is not present in the database is calculated as following

$$V(Z) = \sum_{g \in G} \sum_{i=1}^{N_d} w_{g\alpha}(Z) T_{g\alpha}(Z). \quad (2)$$

In equation (2), g represents the symmetry group of the current configuration. For example, Collins and Bettens classified the BeH_3 system to have five different symmetry groups.³⁷ The advantage of this classification is that it is not required to retrieve all configurations present in the database at every potential calculation. In fact, retrieving partial databases helps to reduce much of the huge computational cost, which is a disadvantage of this fitting technique compared to the other methods. $w_{g\alpha}$ represents the weight of point i . The form of the weight function was discussed in details in some previous works.⁶⁵⁻⁶⁷ By fitting 1,300 data points of BeH_3 system, the absolute average error on a testing set of 1,000 points was reported as 0.41 kJ mol^{-1} .⁶⁷ The fitting errors using Shepard interpolation method vary from system to system. They range from 0.1 kJ mol^{-1} to 3.0 kJ mol^{-1} .

For the first time, MD trajectories are employed to sample configurations in developing *ab initio* PES's as proposed by Collins and co-workers.³³ They have established a detailed summary of the Shepard interpolation method, and presented the PES's of many different molecular systems.⁶⁸ In that paper, they proposed that a set of molecular configurations chosen from the reaction paths had a certain role in describing the reaction dynamics. The data-growing process is executed over and over, until a

convergence criterion is satisfied. This criterion is based on the convergence of the dynamics and therefore, requires repeated computations of the reaction dynamics resulting from the current PES with the ones resulting from a previous PES that fits less data points. If the difference in results is sufficiently small, the system is believed to converge and no further sampling is required.

Although the Shepard interpolation method is a very feasible method, there are still some critical difficulties that prevent it to become a powerful and universal method to many molecular systems. Until today, $H_3^+ + HD$ and $H_3^+ + D_2$ is the most complicated systems that has been investigated using the Shepard interpolation method.⁶⁹ For an N -body system, the number of input coordinates is $3N-6$; therefore, the number of terms in each Taylor expansion extended through quadratic form is $4.5N^2 - 13.5N + 10$. The computational requirements of the Taylor expansion terms rises significantly if we compare the previous cases to a four-body system like BeH_3 to $H_3^+ + HD$. Retrieving data is absolutely required at every new generated configuration during a MD trajectory. The weights must also be re-evaluated (this may be considered as performing a whole new fit for a new configuration). All of these facts lead to a difficulty of computational cost, which is the major concern in any molecular dynamics studies until nowadays despite the fact that computer technology has grown extremely fast in the past few decades. Lastly, the numerical *ab initio* derivative information has to be available in order to perform an interpolation, which makes this method impractical for systems that require high quality electronic structure calculations (for example, in MP4 calculations, numerical gradients are not available for quadruple excitations).

II. Interpolating moving least square (IMLS) method

Interpolating moving least square method has been in practice for several years. The potential energy is expressed as a combination of linear basis functions. The mathematical form is given as:

$$V_{fitted}(Z) = \sum_{i=1}^M a_i(Z)b_i(Z), \quad (3)$$

where M is the number of basis functions, a is the coefficient, b is the basis function vector, and Z is the input vector. Again, in this technique, Z consists of the inverse atomic distances. We can see clearly from equation (3) that both a and b are two functions of Z , and they have to be calculated at every new configuration. The basis functions are to be constructed prior to the MD investigation. At every point during MD calculations, coefficients a_i are computed in order to minimize the sum of weighted square errors of all points in the database. This is done by solving a matrix equation of size ($N \times M$) at every new configuration during the MD investigation, and thus, the computational cost of this technique is huge. The cost of a fitted potential is proportional to NM^2 , where M comes from the number of basis functions, and NM is resulted from solving the matrix equation. Guo *et al.*⁷⁰ have reported the computational time for one H_2CN trajectory lasting for 2.5 ps in molecular time was about 6,000 seconds. This method requires much more computational time than the Shepard interpolation method.

Lowering the computational time effectively has always been a major concern when the IMLS technique is used. Recall that in the Shepard method, data points are classified into different symmetry groups to improve the speed. Here, in this method, a technique known as “cutoff radius” is also used to deal with the problem. The determination of the “weights” and coefficients a_i is made based on some criteria, such as

calculating the distance between a new configuration and a certain configuration in the database. If the distance is not in the range of the cutoff radius, that configuration in the database is considered as not useful in the fit, and interpolating around that point is not required. The choice of the cutoff radius has to be done wisely, as it is a trade-off between computational cost and fitting accuracy.

It has been shown that the fitting error goes down as more data points are used. Once the IMLS method is employed, one has to determine the trade-off between fitting error and computational speed. In the study of HOOH, Guo *et al.*⁶¹ performed three fits with various number of points using the valence internal coordinates (which consist of three bond distances, two bending angles, and a dihedral angle) and three fits with various number of points using the inverse interatomic distance coordinates. The best fitting error (8.25 kJ mol⁻¹) has resulted from the fit using valence internal coordinates with 1,000 data points. In our HOOH study using the NN method, we also observed experimentally that the use of valence internal coordinates gives better accuracy than the use of interatomic distances as input coordinates.

The local IMLS method (L-IMLS) restricts the number of weight evaluations.⁷⁰ For each point in the database, the fit is performed, and the resulting coefficients are stored. Once enough data points are collected, evaluations of the potential at a new point are performed based on the existing coefficients. The L-IMLS method helps to reduce computational cost.

The advantage of this technique over the Shepard interpolation method is the capability to fit data from high quality *ab initio* calculations, most of which do not provide first and second derivative information. Moreover, since first and second

derivatives are not extracted from *ab initio* calculations, IMLS method saves more time in electronic structure calculations.

III. Neural network method

Neural network fitting²² is a computational method that has been applied to chemical reaction dynamics since the 1990s, and broadly used to fit many numerical functions. In the field of MD study, NN has been proven to be powerful and robust because the fit is executed only once and available for use throughout the study. It has been employed to fit many PESs with various levels of complexity. Unlike the Shepard and IMLS methods, NN fitting provides an analytical function, which is very helpful when gradient analysis is executed.

Let us denote vector R as the atomic distance coordinate (or valence internal coordinate) vector. Prior to the fitting process, R and the *ab initio* potential energy V are scaled as following

$$r_{scaled_i} = \frac{2(r_i - r_{i_min})}{r_{i_max} - r_{i_min}} - 1, \quad (4)$$

$$V_{scaled} = \frac{2(V - V_{min})}{V_{max} - V_{min}} - 1, \quad (5)$$

where r_{i_min} and r_{i_max} are the i^{th} minimum and maximum atomic distances, respectively. Similarly, V_{min} and V_{max} are the minimum and maximum potential energies, respectively. This scaling maps all number (inputs and outputs) in the range of -1 to +1.

We employ a two-layer feed-forward NN to fit the function. The output of the first layer is produced as following

$$n_i = \sum_{j=1}^6 w1_{j,i} r_{scaled\ j} + b1_i \text{ for } i = 1..m. \quad (6)$$

Each calculation of n_i in the first layer is defined as a “hidden neuron.” m is the number of hidden neurons, $w1$ is the weight matrix of size (m x 6), and $b1$ is the (m x 1) bias column vector. In all three studies that we have conducted, there are 6 inputs (either valence internal coordinates or atomic distance coordinates).

The outputs of the first layer are converted to be the inputs for the second layer by a transfer function f^d . Hornik *et al.*⁷² have shown feed-forward NNs that use a sigmoid function as the transfer function in the first layer, and a linear function as the transfer function in the second layer are universal approximators for analytic functions. Specifically, in our studies, we use the hyperbolic tangent function as the transfer function in the first layer. The NN output is computed as

$$V_{out} = \sum_{i=1}^m w2_i f^1(n_i) + b \quad (7)$$

In equation (7), $w2$ is the weight row vector of size (1 x m), and b is a scalar number which represents the bias value of the second layer. The number of neurons is assigned by users. The data is divided into three sets: training, validation, and testing. Subsequently, the NN is trained to minimize the sum square error of the training set using *Levenberg–Marquardt* algorithm.²² Usually, the training set consists of 80% of the data, while each one of the other two sets consists of 10%.

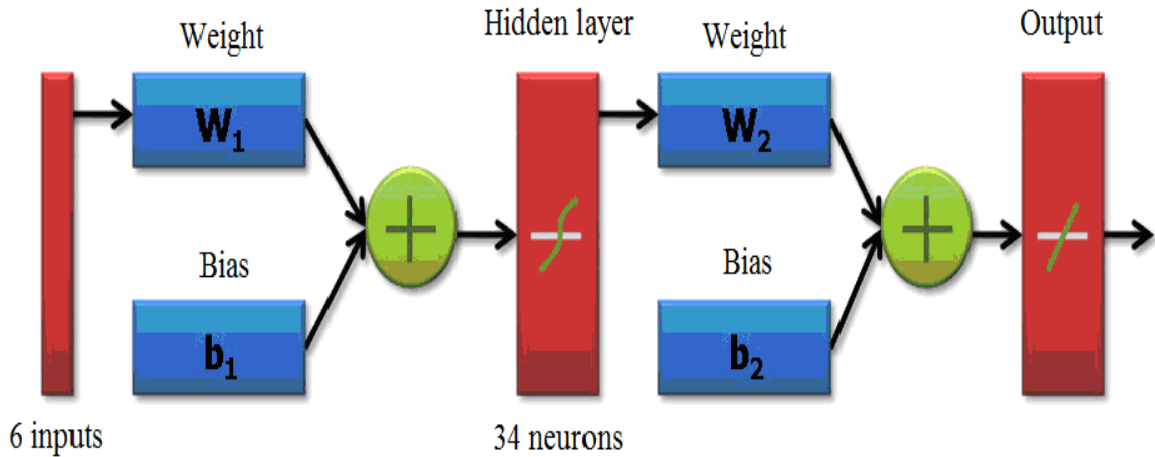


Figure II.1 Structure of a two-layer feed-forward NN with 34 neurons in the hidden layer

Over-fitting is a very serious issue that is of concern when using the NN method.⁷³ An illustrative example of over-fitting is given in Figure II.2. There are several ways to avoid this problem, but the most common technique is known as “early stopping.”²² A NN is trained for hundreds of epochs (iterations); however, when the sum square error of the validation set is seen to increase in 6 consecutive epochs, the training process is terminated to prevent over-fitting. The weights and biases used are those present when the validation error starts to increase. Figures II.3 and II.4 represent the training process. After each epoch, the root mean squared (rms) errors of three sets are calculated.

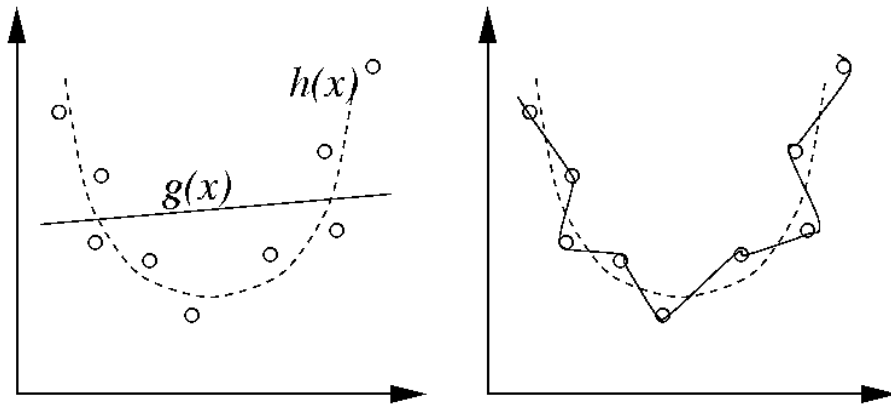


Figure II.2 An illustrative example of over-fitting.⁷⁴

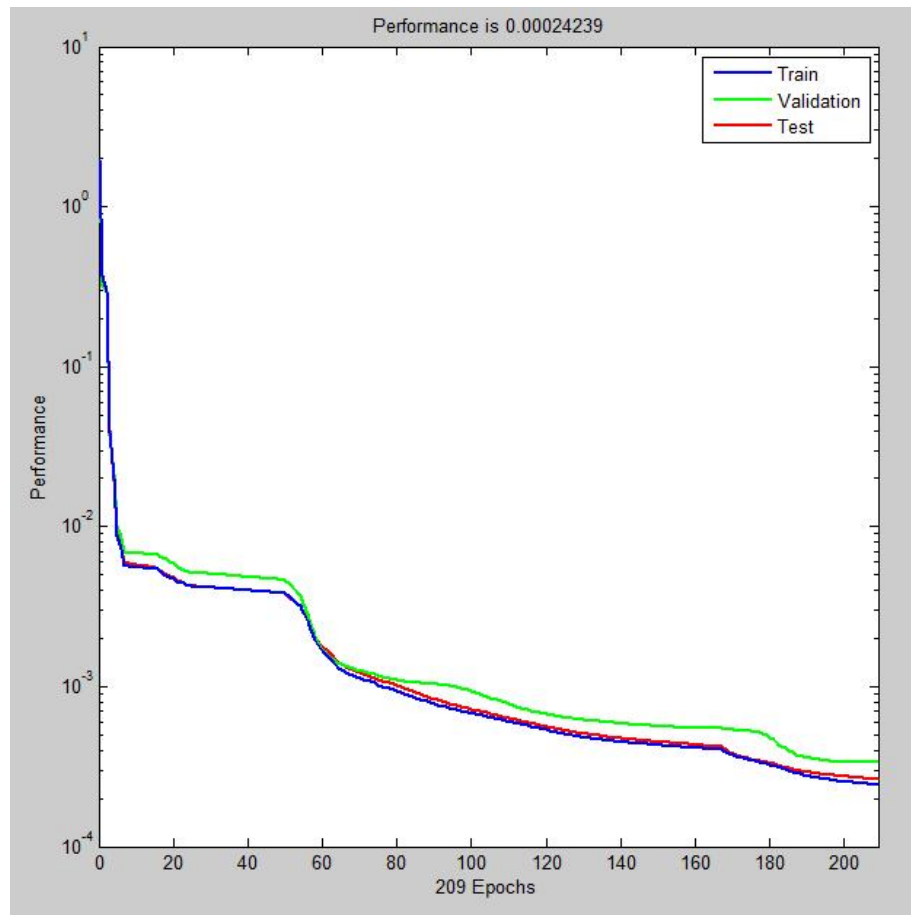


Figure II.3 Mean square error of training, validation, and testing sets.

The training process is terminated after 209 iterations.

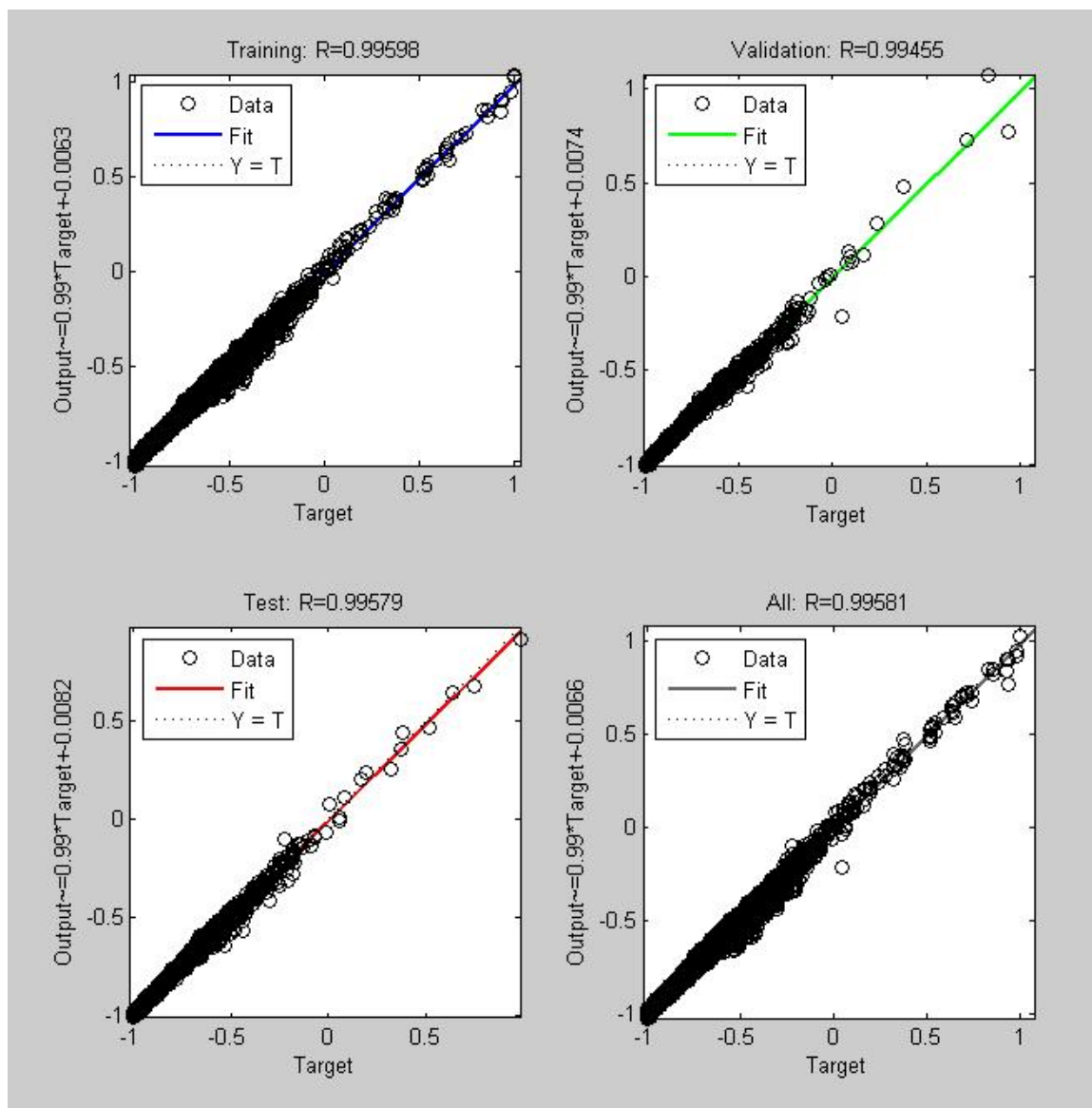


Figure II.4 Plots of outputs vs. targets for training, validation, and testing sets.

The NN method has been employed in many PES studies. It was first employed to study the surface diffusion of CO/Ni(111) system.⁷⁵ To date, several PESs have been reported using this fitting technique. Some of them are CH_2CHBr ,^{29,30,76-79} HONO ,^{24,80} BeH_3 ,³² HOOH ,⁶³ SiO_2 ,⁸¹ hydrogen dissociation on metal surfaces,⁸² reaction of C_2 dimers with an activated diamond (100) surface,⁸³ CO chemisorption on a Ni(111)

surface and the interaction of H_2 with a Si(100)-2x1 surface,^{84,85} and H_3^+ .⁸⁶ Among these, vinyl bromide has been so far the most complicated molecular system studied by any fitting method. More recently, a newly-developed training algorithm was introduced, in which *ab initio* potential energy and the gradients with respect to inputs are fitted simultaneously. Consequently, the method is denoted as a combined function derivative approximation (CFDA) method. This turns out to be a powerful method that describes functions with excellent accuracy. The method was applied to Si_3 , Si_5 clusters,⁸⁷ and the H+HBr system.⁸⁸

Low computational cost is a big advantage of the NN method. Since the PES fit is done only once, it is very convenient for users to extract analytic derivatives. More data points would require more neurons to perform the fit; however, it does not increase the number of neurons significantly, and can be considered as a negligible issue. In our three studies, we obtain about 20 thousands points for each case, and the fitting errors in all cases turned out to be very low.

Because of low computational cost, we generally use a NN committee to enhance fitting accuracy. Several NN's are trained individually by randomly selecting training data. Subsequently, the NN's committee is taken to be an average of all individual NN's. As will be shown latter, the error of NN committee is less than the errors resulted from each individual NN as a result of nearly random error cancellation.

IV. Support vector machine method

A support vector machine is a "cousin" tool of NN fitting, which has been proposed and developed after the birth of NN's. This technique is one of the most common machine learning algorithms and has been used for a variety of chemistry

applications, such as optimization of chromatographic separation, concentration prediction from spectral data, qualitative and quantitative prediction from sensor data, and so on. An overview of SVM applications in chemistry is available in ref. 89 for consulting. The SVM algorithm was initially developed by Vapnik and Lerner⁹⁰ based mainly on the statistical learning theory. Drucker *et al.*⁹¹ latter developed the fitting algorithms for regression training as one important feature of SVM method (also known as support vector regression, SVR) to fit analytic function.

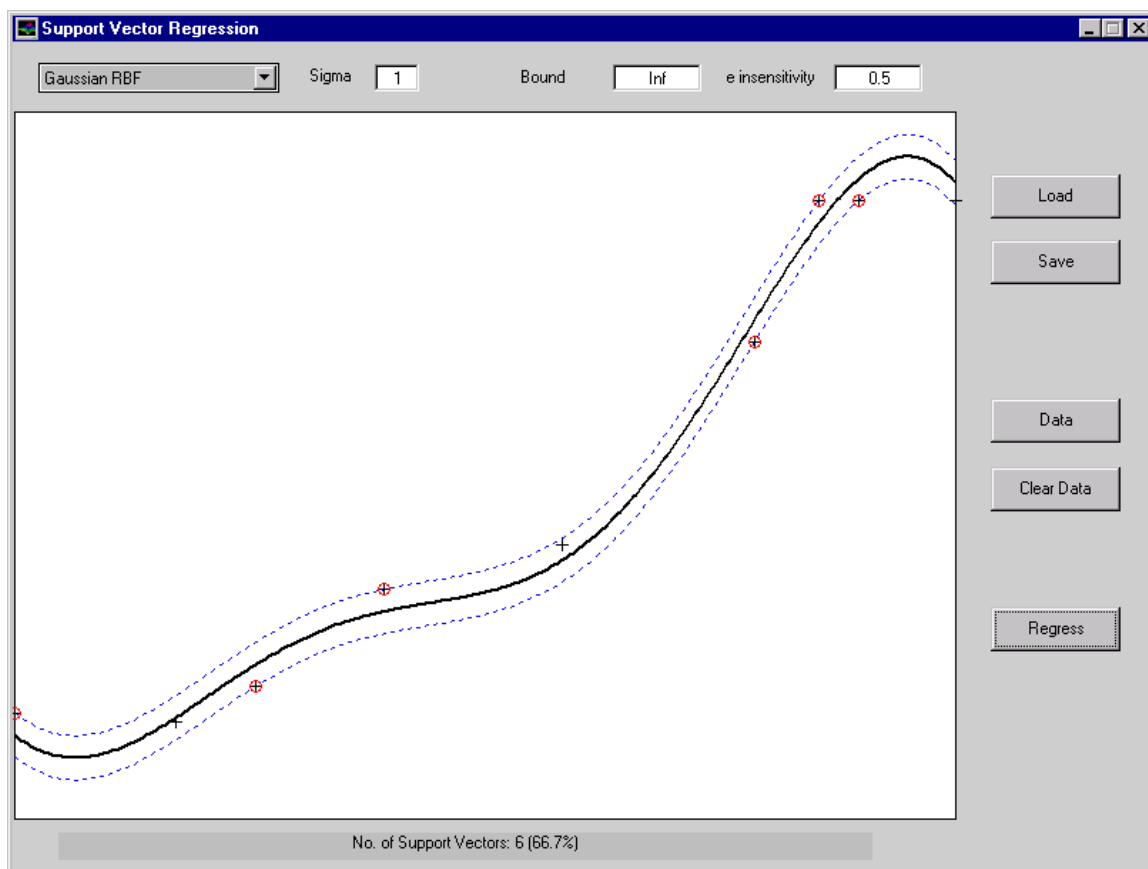


Figure II.5 Example of an SVM regression using radial basis kernel.⁹²

Source: <http://users.ecs.soton.ac.uk/srg/publications/pdf/SVM.pdf>

The success of the NN method in MD studies has led us to explore the ability of an SVM model to fit *ab initio* data. Similarly to a NN, an SVM fit is an analytic function that only needs to be trained once. Prior to conducting the SVM *ab initio* PES study, our expectation is that SVM method will be comparable to NN method in terms of computational cost and accuracy.

We also use equations (4) and (5) to scale inputs and output in the region of -1 to +1. The mathematical form of an SVM function using the radial basis kernel is

$$V_{out}(x) = \sum_{i=1}^n \alpha_i e^{-\gamma(v_i-x) \cdot (v_i-x)} - \rho \quad (8)$$

where γ , ρ , and α_i are scalar values (γ is defined by users), n is the number of support vectors required for fitting with some given criteria, v_i is a vector of six dimensions, and x is the (6 x 1) input vector. Each term in the above summation is referred to a “support vector.”

There are also some other important parameters that need to be defined by users prior to training. The “cost” c is a very important parameter which determines the training speed and number of support vectors. Technically, c represents the absolute values of α_i in the function. A high value of c gives a wide range of α , which makes the effective region of a support vector larger, and thus, would decrease the number of support vectors. However, it is very time consuming to train with high value of c . The epsilon value of the “loss function” ρ has a direct effect on accuracy. Obviously, smaller value of ρ gives better accuracy, but costs more computer time to train the data. The estimation of γ , c , and ρ will be discussed latter as we attempt to fit the HOOH data using SVM method. Some comparisons of PES’s obtained using both NN and SVM methods will be made in a subsequent chapter.

CHAPTER III

CIS-TRANS ISOMERIZATIONS AND N-O BOND DISSOCIATION OF NITROUS ACID (HONO)

I. Development of *ab initio* potential surface

1. Electronic structure calculations

The *ab initio* potential energy calculations are executed using the Gaussian suite of programs.⁹³ In this work, accuracy is given the first priority and computational cost is our second consideration. Therefore, we choose to employ the fourth-order Moller-Plesset perturbation method⁴⁴ [MP4(SDQ)] with singlet, doublet, and quartet excitations. Other methods have also been executed for some comparisons, such as MP2 and couple-cluster method^{94,95} (CCD) with double excitations.

The choice of basis sets does not have a big affect on computational cost; in fact, accuracy and convergence are two direct consequences of a basis set choice. Several basis sets have been tested and compared to find the most suitable one. They include 4-31G,⁹⁶⁻⁹⁹ 6-31G, 6-31G(d),¹⁰⁰ and 6-311G(d).^{101,102} Most *ab initio* calculations in the sensitive region that describes N-O dissociation (large N-O distance) on 6-31G and 6-31G(d) basis sets fail to converge. The 6-311G(d) basis set, on the other hand, has almost no problems in the convergence issue when MP4(SDQ) level of theory is applied. Also, we observe that the potential energy obtained from MP4(SDQ)/6-311G(d) is lower than

results from the other methods. Therefore, it is chosen for the *ab initio* calculations in this study. The couple-cluster method with double excitation (CCD) on the large triple zeta basis set⁹⁻¹³ (cc-pVTZ) is also executed, but this method is not preferred in the study because of huge computational cost compared to the others.

The potential barriers of *cis-trans* isomerization and N-O dissociation are obtained from partial optimization of the potential energy using the Newton-Raphson method. In a partial optimization, some Gaussian input parameters are set as constants, and the rest are optimized to minimize the potential energy. To investigate the rotational barrier, several partial optimizations are executed for various dihedral angles from 0 to π , and the corresponded potential energies are recorded.

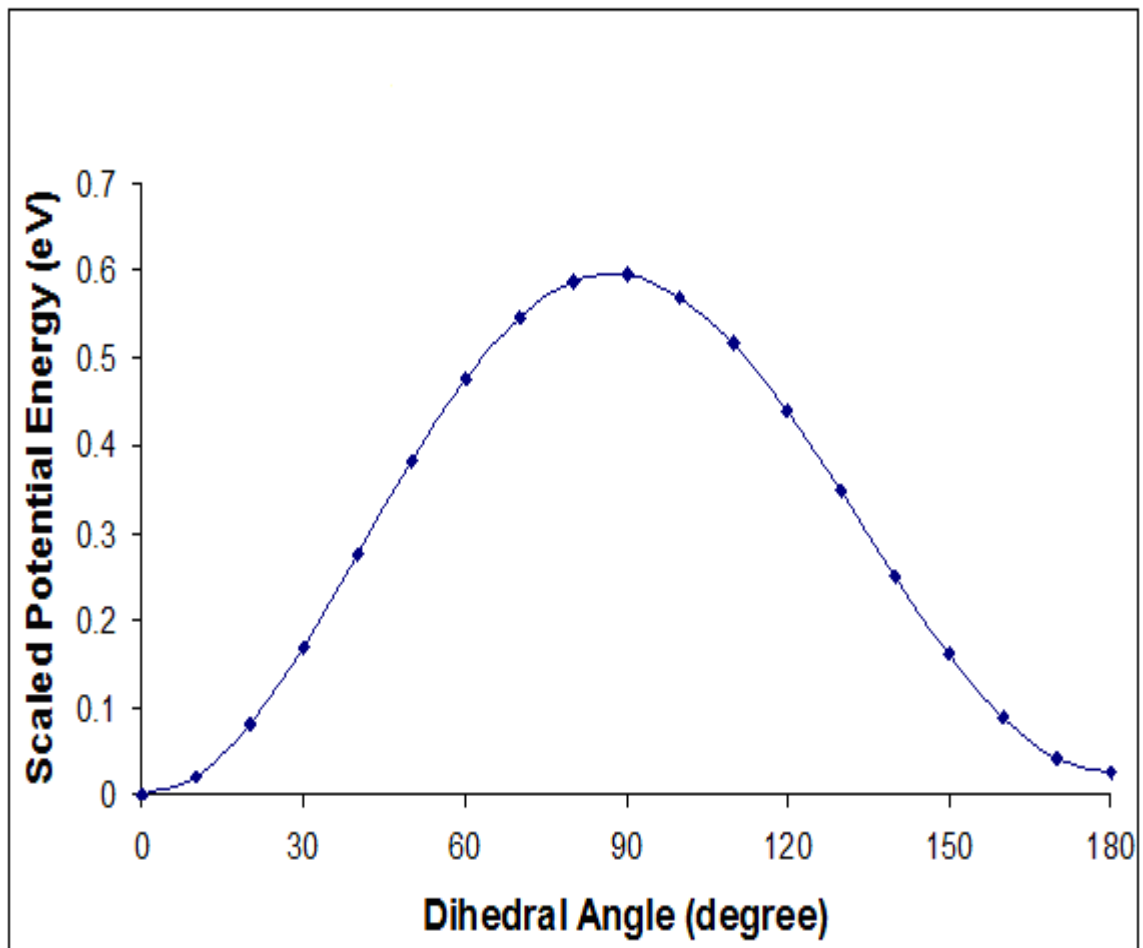


Figure III.1 Rotational barrier of HONO

Similarly, the N-O dissociation barrier is investigated by optimizing the potential for various N-O distances. From the potential barrier, it is concluded that the dissociating distance is 2.78 Å.

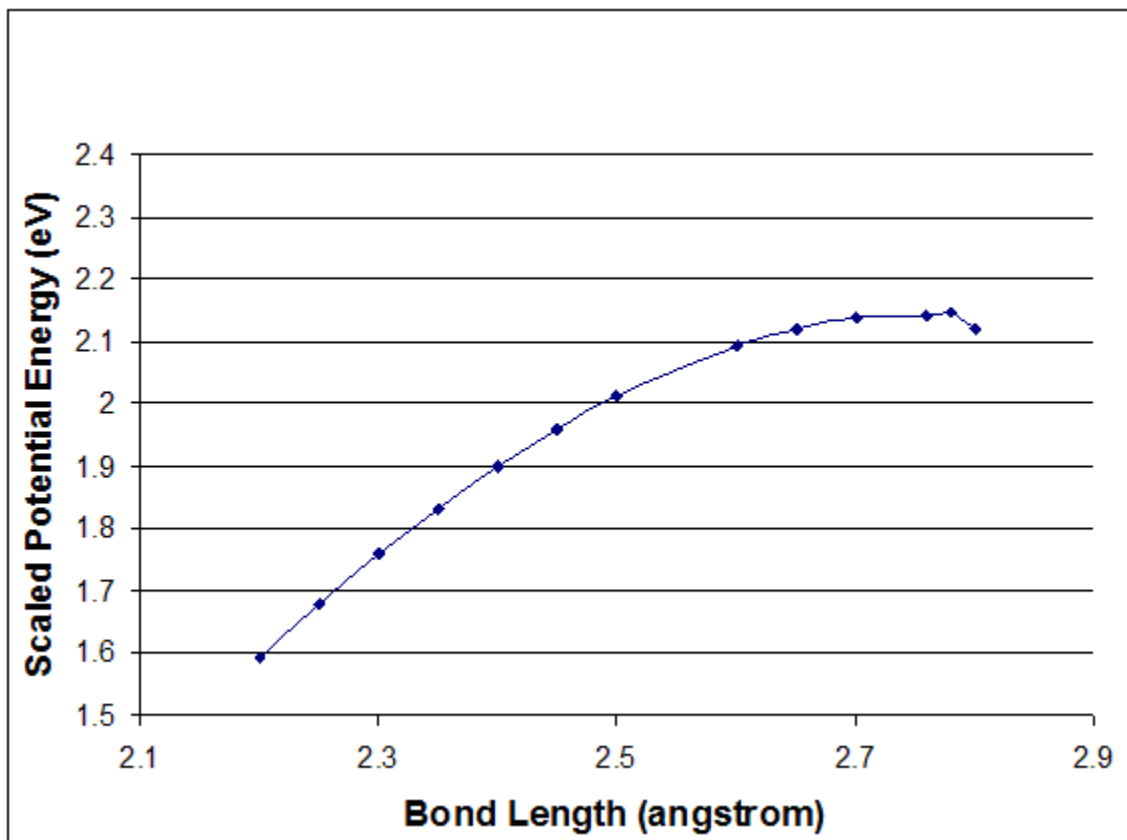


Figure III.2 N-O dissociation barrier of HONO

From MP4(SDQ)/6-311G(d) calculation, the rotational barrier is 0.552 eV, while the dissociation barrier is 2.15 eV. When we compare the results among several calculations, we find that MP2 level of theory results in highest rotational barriers, while CCD/cc-pVTZ gives the highest dissociation energy (Table III.1).

Table III.1 Equilibrium energies and potential barrier heights for various computational methods and basis sets

Ab initio method	<i>cis</i> equilibrium energy (Hartree)	Rotational barrier (eV)	N-O dissociation barrier (eV)
MP2/6-311G(d)	-205.2618450	0.595	2.54
MP2/6-31G*	-205.1671423	0.607	2.46
MP4(SDQ)/6-311G(d)	-205.2675897	0.552	2.15
MP4(SDQ)/6-31G*	-205.1747915	0.565	2.12
CCD/cc-pVTZ	-205.2022592	0.515	3.16

By performing two full optimizations of HONO, in which all parameters are considered as variables, it is found that there are two energetic minima. The *cis* configuration is more energetically stable than the *trans* configuration.

A calculation for one HONO configuration operated on a processor with a 2.4-GHz-clock-speed using MP4 level requires about 45-50 s to accomplish, which mean for a set of 21,584 configurations, about 280 hours of computer time are needed to execute *ab initio* calculations to obtain the potential energies.

2. Modified novelty sampling procedure

According to the novelty sampling procedure developed by Raff and co-workers,³¹ a temporary potential surface must be available, and more configurations are obtained by executing MD trajectories on this temporary potential surface. *Ab initio* calculations are executed for the new configurations, and they are added to the database only if they satisfy the novelty sampling selection criteria. The more points we select, the more converged the PES becomes. However, some criteria have to be used in order to

test for convergence and terminate this sampling process, as we do not want to collect too many unnecessary data points. After adding the new points, the percentage of selected points to the current database is calculated. If the number of added data points is small, we believe the database has converged. The novelty sampling procedure has been found to work effectively in the case of vinyl bromide.^{29,30}

In our study of HONO, we employed a simple empirical surface developed by Guan and Thompson³ to serve as the first temporary PES. Several trajectories, each of which is initialized with a random geometric configuration and random momenta using the projection method (will be discussed in the trajectory calculation section), are integrated for 5 ps with a fixed step size of 3 (as 0.01 time unit) with the internal energy of 2.8 eV (including zero point energy), and the Cartesian coordinates are recorded. In the first iteration, we obtain 4,000 configurations with trajectories starting from *cis* configurations, and 4,000 configurations with trajectories starting from *trans* configurations, which combine to make a first set of 8,000 data points. The *ab initio* potential energy for each configuration is then calculated using MP4(SDQ)/6-311G(d) level of theory. Six atomic distances are recorded as inputs, and *ab initio* potential energies are recorded as outputs.

A temporary NN surface is constructed based on 8,000 data points. More configurations are generated from trajectories on this temporary surface. We have developed some criteria to examine the configurations before adding them into the database. Let us denote vector v_i as a seven-component vector, which consists of six normalized inputs and one NN output which corresponds to the six inputs. The seven

components of vector v_i are all unitless. The distance between the i^{th} vector and the j^{th} vector is defined as

$$d_{ij} = \sqrt{\sum_{k=1}^7 (v_i - v_j)^2} . \quad (9)$$

The i^{th} minimum and average distances are directly obtained from the calculated values of d_{ij} .

$$\min_i = \min(d_{ij}) \text{ where } j = 1 \dots n, \text{ and } n \text{ is the total number of vectors,} \quad (10)$$

$$\text{avg}_i = \frac{\sum_{j=1}^n d_{ij}}{n} . \quad (11)$$

$P(\min(v_i))$ and $P(\text{avg}(v_i))$ are the probability distribution functions of minimum and average distance, respectively. There are two designated tests to examine a new configuration. A generated configuration from NN trajectories is qualified to the next progress if it passes one of the following

$$\text{if } P(\min(v_k)) \leq T_1 P(\min(v))_{\max} \text{ and } \min(v_k) \geq \alpha_{\min} , \quad (12)$$

$$\text{or if } P(\text{avg}(v_k)) \leq T_2 P(\text{avg}(v))_{\max} \text{ and } \text{avg}(v_k) \geq \alpha_{\text{avg}} , \quad (13)$$

where T_1 and T_2 are selected as 0.5 and 0.2, respectively. From the probability distribution plots of minimum and average distances, we also choose α_{\min} and α_{avg} as 1.2 and 0.1, respectively. The purpose of condition (12) is to ensure that if a point is far away from the other configurations in the database, it is a useful point. However, if a point is close to some other in the database (which means it would fail condition (12)), and if that group is far away from the majority, the point is still considered useful, and condition (13) helps to qualify that point.

An effective microcanonical method to sample the configuration hyperspace for some molecular systems, including HONO, was discussed by Schranz, Nordholm, and Nyman.¹⁰³ This procedure, however, includes many configurations that are not useful to the reaction dynamics of HONO. Therefore, we employ a modified novelty sampling procedure³¹ to acquire data. MD trajectories usually produce more data points in the near-equilibrium region, and less in the reacting region. However, the reacting region is more important in constructing the PES. Therefore, it is important to recognize those well-described points and eliminate the unnecessary ones from the data. Empirically, it is observed that adding more data points to the regions of configuration space that is adequately covered would increase the fitting error in the other important regions. Therefore, a direct comparison of *ab initio* energy and NN predicted energy is made, if the percent difference is less than 1%, the point is discarded; otherwise, it is selected.

After selecting a set of data points, Gaussian input parameters are calculated, and *ab initio* calculations are performed. The resulting *ab initio* potential is compared to the NN potential as a final test. This process is executed iteratively until convergence, in which we are not able to select many additional data points.

With an initial set of 8,000 data points, a temporary surface is trained. In the first iteration, we mainly focus on finding data near the equilibrium region. Some trajectories are started at either *cis* or *trans* equilibrium configuration with no kinetic energy, and integrated for a period of time. We sample those configurations, perform *ab initio* calculations and add them to the database. During the course of this first iteration, 1,000 more points are added to the database.

Subsequently, four more iterations are executed to add more data points into the system. In the last four iterations, bond stretching is our major goal in obtaining data points, especially N-O dissociation. Various levels of internal energy are applied to ensure that the important regions are adequately covered. As shown in Figure III.3, 3,314 data points are added to the system in the second iteration. Most added points lie in the high minimum-distance region, which we do not see in the previous database.

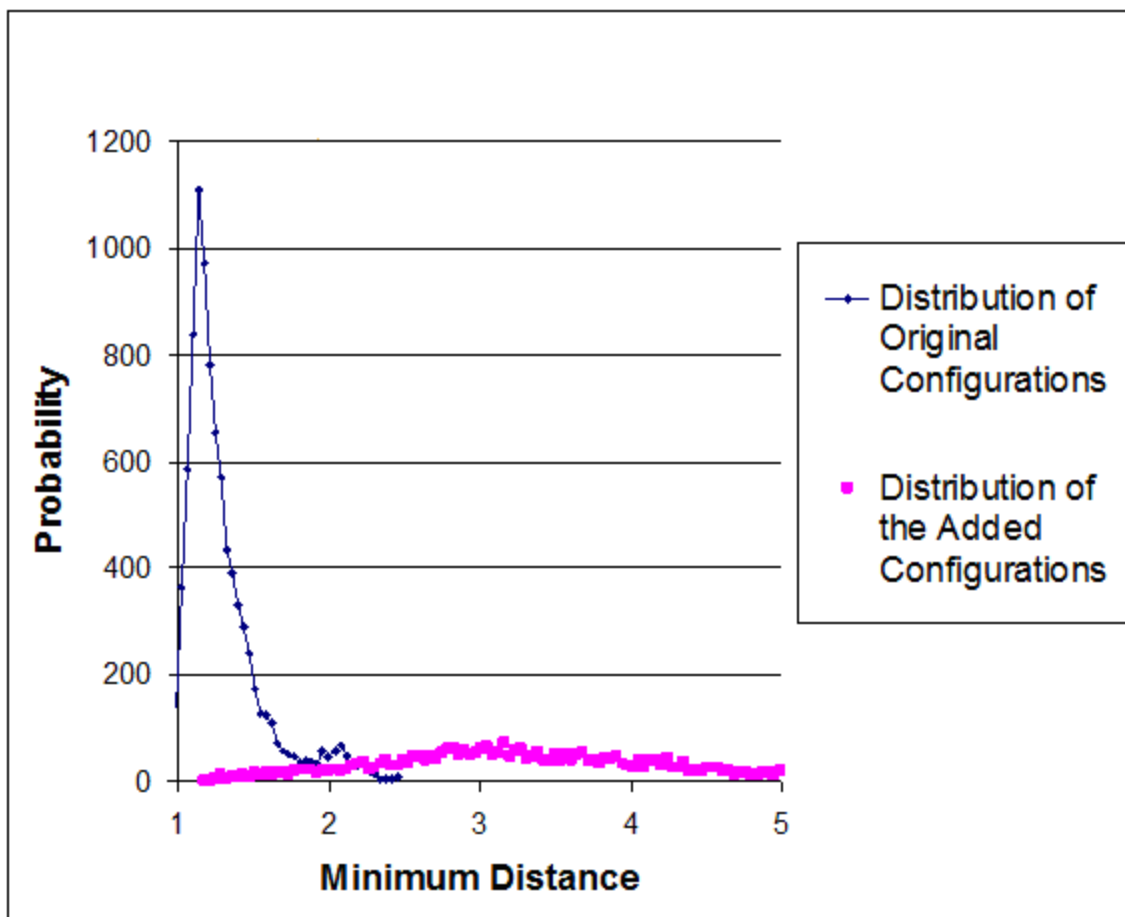


Figure III.3 $P(\min(v_i))$ vs. minimum distance. Most of the points from the overlapped region are rejected. 3,314 more data points are added to the system in the second iteration, which mainly focus on the stretching of three chemical bonds.

During the sampling process, we have noticed a circumstance when N-O bond is stretched too far apart. At a large N-O distance, MP4(SDQ) method fails to converge, and as a result, it either terminates the calculation, or gives non-sense potential energy (much lower than the equilibrium energy). After five iterations, we have selected 21,584 configurations. The minimum distance distribution for the last iteration is shown in Figure III.4. The detailed number of added configurations during each iteration is given in Table III.2.

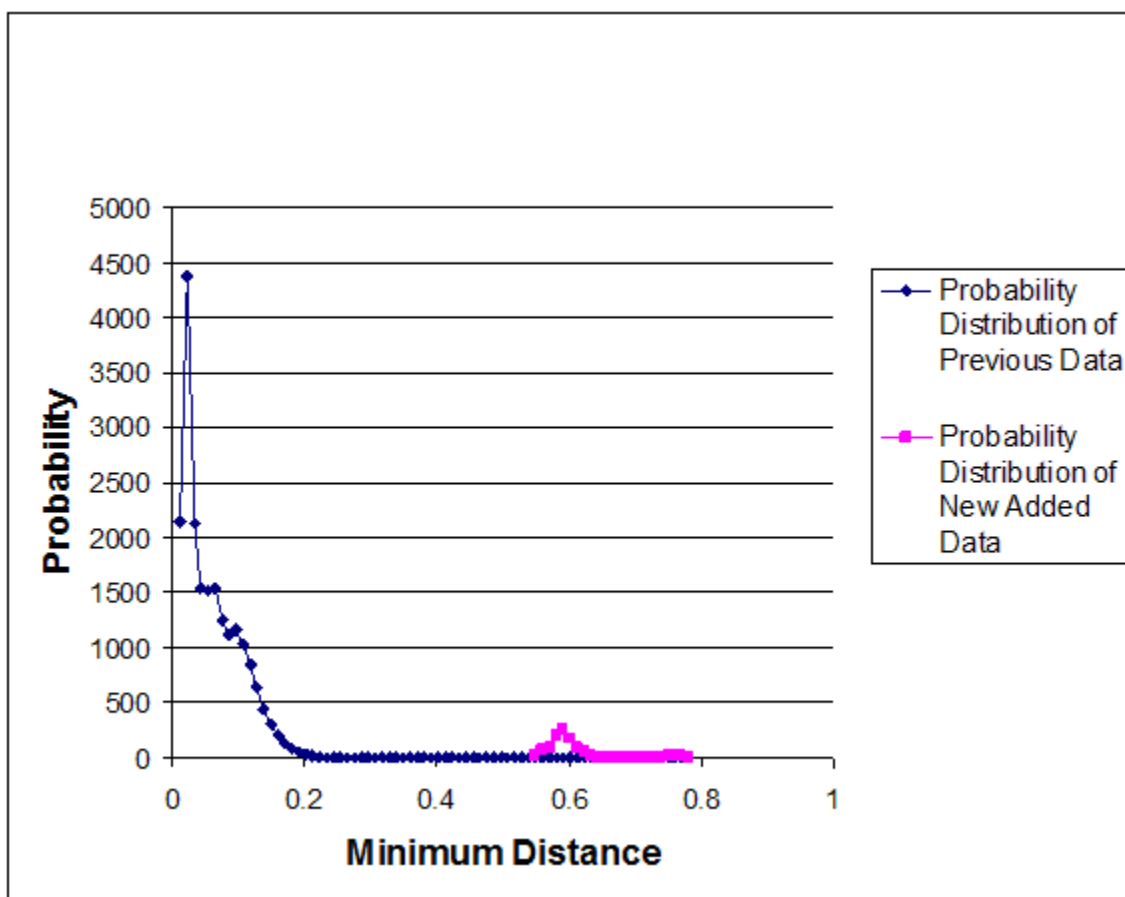


Figure III.4 Minimum distance distribution for iteration 5. As we notice, most of the new data overlap with the previous data, and not many configurations can be selected. At this time, we decide to terminate the sampling process.

Table III.2 Number of added configurations in each iteration process

Step	Starting	Adding	Total	Note
0	0	8000	8000	Initial configurations
1	8000	1000	9000	Focusing on <i>cis</i> and <i>trans</i> equilibrium configuration
2	9000	3314	12314	Bond stretching
3	12314	6288	18602	Bond stretching
4	18602	1996	20598	More data described O-N dissociation
5	20598	986	21584	More data described O-N dissociation

Table III.3 Equilibrium, maximum and minimum atomic distances (Å) of HONO

Identity of the atomic distance	O ¹ -N (r ₁)	N-O ² (r ₂)	O ² -H (r ₃)	O ¹ -O ² (r ₄)	O ¹ -H (r ₅)	N-H (r ₆)
<i>Cis</i> equilibrium	1.195	1.380	0.969	2.161	2.119	1.901
Exp. <i>cis</i> equilibrium	1.185	1.392	0.982	2.159	2.108	1.877
<i>Trans</i> equilibrium	1.179	1.381	0.960	2.102	2.896	1.881
Exp. <i>trans</i> equilibrium	1.170	1.432	0.958	2.146	2.883	1.882
Maximum	1.527	2.850	1.391	3.873	4.881	3.812
Minimum	0.976	1.079	0.775	1.673	1.381	1.330

The six equilibrium atomic distances of *cis* and *trans* configurations are given in Table III.3. We have found that the result is in good agreement with the experimental values.¹⁰⁴

3. Normal mode analysis on the new PES

Vibrational wave numbers are an important characteristic of Born-Oppenheimer potential energy surfaces, in which we employ some mathematical approximations to extract the wave numbers. The results from this procedure can be compared directly to the wave numbers calculated by the corresponding *ab initio* method (MP4(SDQ)/6-311G(d) in the case of HONO), and fitting quality can be evaluated.

The details of normal mode analysis have been described by Herzberg¹⁰⁵ and Wilson, Decius, and Cross.¹⁰⁶ Consider a 2-dimensional matrix D of size (3N, 3N) for a molecule (where N is the number of atoms), where the element D_{ij} is calculated as

$$D_{ij} = \frac{1}{\sqrt{m_i m_j}} \frac{\partial^2 V}{\partial q_i \partial q_j}, \quad (14)$$

where m_i is the atomic mass corresponding to coordinate i .

As we can see, D is a Hermitian matrix, which always yields 3N real positive eigenvalues.

$$D = \lambda L. \quad (15)$$

L is an orthonormal matrix, which consists of 3N eigenvectors, λ is a diagonal matrix, each value of the diagonal represents the fundamental frequency. If the molecule is non-linear, 6 frequencies, which represent translational and rotational modes, are expected to be very small (close to 0), and the other 3N-6 values are the vibrational frequencies. If the molecule is linear, there are 3N-5 vibrations. L is a useful orthonormal matrix, which can be employed to distribute energy into vibrational modes.

In the HONO case, there are two global minima associated with the *cis* and *trans* configurations. Normal mode analysis is executed at both conformations, and the results shown in Table III.4 are in excellent agreement with *ab initio* values. They are higher

than the experimental values because the fitting is done to the MP2 results. Recall that this procedure is based on the second derivative analysis, which means the NN fits the shape of the function very well. This is a result of the sampling method which accurately fits the surface gradients.

Table III.4 Experimental and calculated wave numbers (cm^{-1}) of *cis* and *trans* HONO

Vibrational Modes		OH	N=O	HON	O-N	ONO	Torsion
<i>Cis</i>	Experimental ¹⁰⁷	3426	1641	1302	852	609	640
	Calculated by NN surface	3597	1639	1370	968	671	710
	Gaussian (MP2)	3608	1643	1382	958	748	671
<i>Trans</i>	Experimental ¹⁰⁷	3591	1700	1263	790	596	544
	Calculated by NN surface	3799	1738	1348	967	708	549
	Gaussian (MP2)	3796	1694	1338	868	645	616

4. Neural network committee

It has been shown that using NN in MD studies is a time saving method, since the fit is done only once. When higher fitting accuracy is required, NN methods provide two possibilities. First, increasing the number of hidden neurons is an obvious option. We have performed several fits by varying the number of hidden neurons, and learned that a NN with 41 neurons is sufficiently effective for the HONO surface. Therefore, 41 neurons are used to produce a good fitting error of 0.017 eV. Using more hidden neurons may result in over-fitting NNs. To improve the fitting accuracy, our second option is to construct a NN committee, which is a combination of several NN's, each having 41 neurons in the hidden layer.

Different training data selection results in slightly different NN parameters. Selecting data randomly for the training set can be employed to construct different NNs. Also, one can vary the percentage of training data, which also results in a slightly different NN. Once several NNs are constructed, the output of a NN committee is taken to be the average of those NN outputs of the members of the committee. Since the data is randomly selected when constructing a network, the fitting errors produced by each member in the committee tend to be randomly distributed about the target output. As we take the average, the nearly random errors are canceled out, thereby makes the accuracy of the committee better than any individual member. If the errors are well randomized, we expect the NN committee error to be reduced by a factor of $N^{-1/2}$, where N is the number of individual NN's.

An obvious disadvantage of using NN committee is the additional computational effort required. At every integration step, gradients have to be analyzed on each individual network, and the averages are obtained. If there are N networks, N different gradient calculations have to be performed, and thus, raises the computational time by a factor of N.

As an illustrative example, we use 5 individual networks to construct a NN committee. The fitting errors of the five networks tested on 21,584 data points vary from 0.0126 to 0.0170 eV. When the NN committee is applied, the fitting error drops to 0.0111 eV. It can be concluded clearly that the performance of the committee is better than any individual network as most nearly random errors resulted from the statistical fitting algorithm are cancelled.

Table III.5 Different (6-41-1) NN Potential Surfaces and Their Coressponding
Average Absolute Errors

NN Potential Surface	Portion of Data Sets			Average Absolute Error (eV)
	Training	Validation	Testing	
Surface 1	0.80	0.10	0.10	0.0170
Surface 2	0.74	0.13	0.13	0.0136
Surface 3	0.70	0.15	0.15	0.0141
Surface 4	0.68	0.16	0.16	0.0126
Surface 5	0.66	0.17	0.17	0.0162
Committee Potential Surface				0.0111

II. Trajectory calculations

This initial sampling technique is described by Raff in a previous application to 1,2-difluoroethane.¹⁰⁸ Initially, all four atoms are assigned at the equilibrium position. Let us denote the Cartesian coordinates to be q_i and the normal coordinates to be Q_i . Matrix L obtained from normal mode analysis is used to transform Cartesian coordinates (or velocities) to normal coordinates (or velocities), as shown in the following equation

$$Lq^{\bullet} = Q^{\bullet} \quad (16)$$

where q^{\bullet} is a 12-component vector representing the 12 Cartesian velocities, and Q^{\bullet} represents 12 normal velocities. The first six components of Q^{\bullet} are the six vibrational velocities; the latter six components are rotational and translational velocities.

Matrix L is orthonormal; thus, there exists a matrix L^{-1} which is the inverse of L , which transforms Q^{\bullet} to q^{\bullet} .

$$L^{-1}Q^{\bullet} = L^{-1}Lq^{\bullet} = q^{\bullet} \quad (17)$$

The first six components of Q^{\bullet} are obtained from converting the experimental frequencies shown in Table III.4 to normal mode velocities. The latter six components representing translational and rotational velocities are set to 0, as we require the total angular momentum to be zero and there is no center of mass motion. After assigning Q^{\bullet} , q^{\bullet} is calculated from equation (17).

With initial Cartesian coordinates and velocities assigned, we integrate the system of partial differential equations on the NN PES for a random period chosen between 0.11 ps and 0.16 ps (two to three times of the longest vibrational period) with a fixed step size of 1.018×10^{-16} seconds using the fourth-order Runge-Kutta method.¹⁰⁹ The total energy is conserved to at least four digits. This process randomizes the vibrational phase of the molecule. After this integration, Cartesian velocities are converted back to normal velocities using equation (16). Excitation energy is then introduced to the desired modes as kinetic energy. Finally, the normal coordinates are converted back to Cartesian coordinates using equation (17). After these operations, the molecule is now guaranteed to have randomized initial vibrational phase angles with zero linear and angular momenta. The trajectory starts from this point, reaction channels as well as any necessary parameters are recorded. The average computational time required to integrate a trajectory for 5 ps (in molecular time) is 4.8 s when a single CPU of 2.4 GHz clock speed is fully utilized.

III. *Cis-trans* isomerizations

Mode specificity of *cis-trans* isomerizations has been investigated in both rare gas matrices²⁵ and in the gas phase.^{3,4} In these two cases, the reaction orders were found to be second and first order, respectively. All theoretical investigations of HONO reported to date have been executed on empirical surfaces. With a total energy of 1.7 eV (including zero point energy), mode specificity was clearly observed.³ The rate of *cis*→*trans* is higher than the rate of *trans*→*cis* in all cases when various vibrational modes are excited. Depending on the excitation, the factors were found to lie among 2.25 to 22.6, which give us a very good picture of mode specificity on the empirical surfaces being studied. In *cis*→*trans* isomerization, the factor of maximum/minimum rate coefficients is 23.8. On the other hand, the factor of *trans*→*cis* isomerization is 3.7.

In the study of HONO, we employ surface 1 in Table III.5 to conduct MD investigation of *cis-trans* isomerizations. As we have investigated, zero point energy of HONO is 0.525 eV, and the potential barrier required to change the conformation (*cis-trans*) is 0.6 eV (13.8 kJ mol⁻¹) according to MP4(SDQ) calculations. It was reported by Guan and Thompson³ that the potential barrier is 13 kJ mol⁻¹,³ somewhat lower than the value reported in our investigation.

In each investigation, the total energy (including zero-point energy) is brought up to 1.70 eV. Six different investigations are executed with various types of excitations. In the first five investigations, zero-point energy is introduced initially, then the excitation energy is introduced into one of the five vibrational modes (excluding the torsional mode). During a previous investigation conducted by Richter *et al.*,¹¹⁰ it was found that the excitation of the torsional mode resulted the fastest isomerization. In the last

investigation, excitation energy is introduced equally among the modes after initiating with zero-point energy.

The monitoring process is executed after the excitation. Taking advantage of the low computational cost of integrating trajectories on the NN surface, 2,000 trajectories with randomized configurations are investigated. Each trajectory is terminated when 5 ps elapses or the change of conformation is found (the dihedral angles reaches 90°). Six first-order decay plots are made from the first order rate law (as shown in equation (18)), and the rate coefficients as well as the statistical standard deviations are obtained directly from the decay curve, as illustrated in Figures III.5 and III.6. The resulting standard deviations are about 2% of the rate coefficients, which do not include Monte Carlo statistical errors. A first-order decay is described as following:

$$\ln\left(\frac{N_t}{N_0}\right) = -kt \quad (18)$$

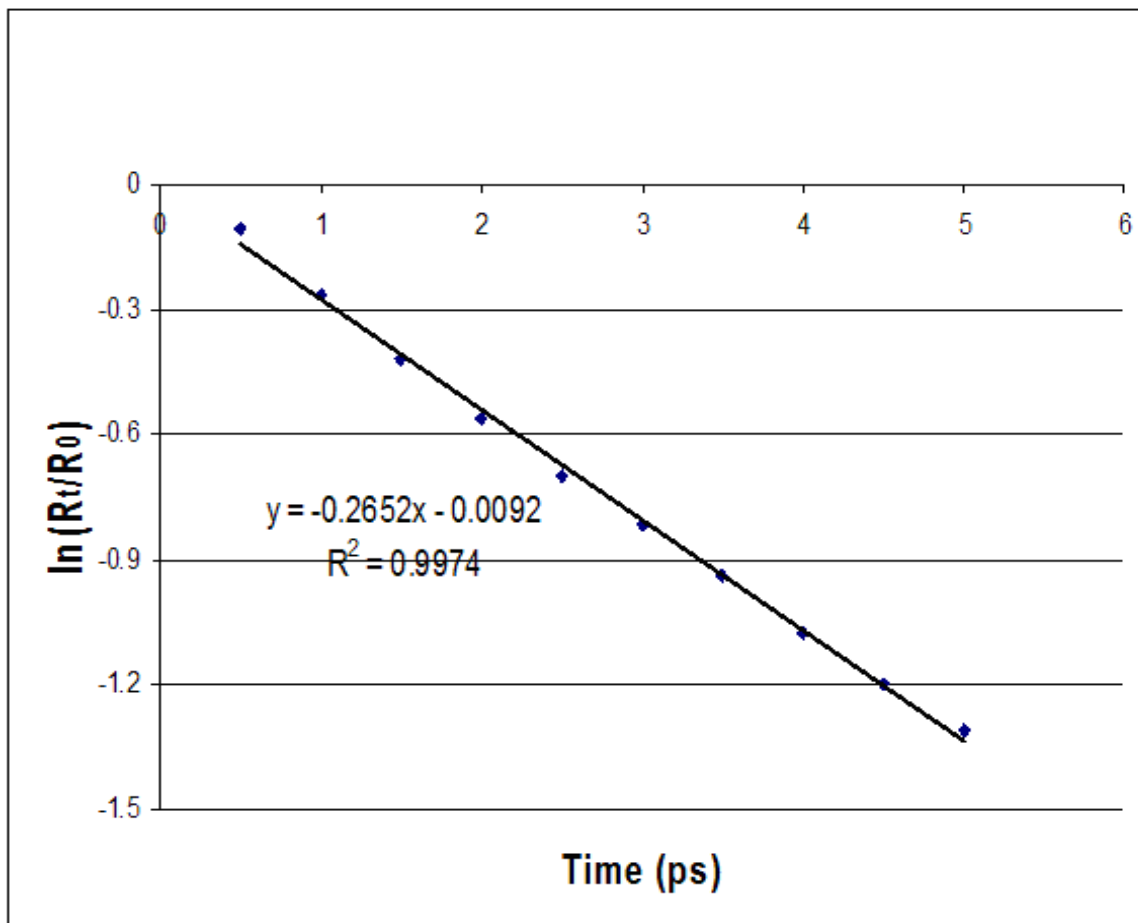


Figure III.5 First-order decay plot of *cis*→*trans* isomerization when OH mode is excited (the internal energy is 1.70 eV, including zero point energy). The curve gives an excellent linearity and statistical standard deviation.

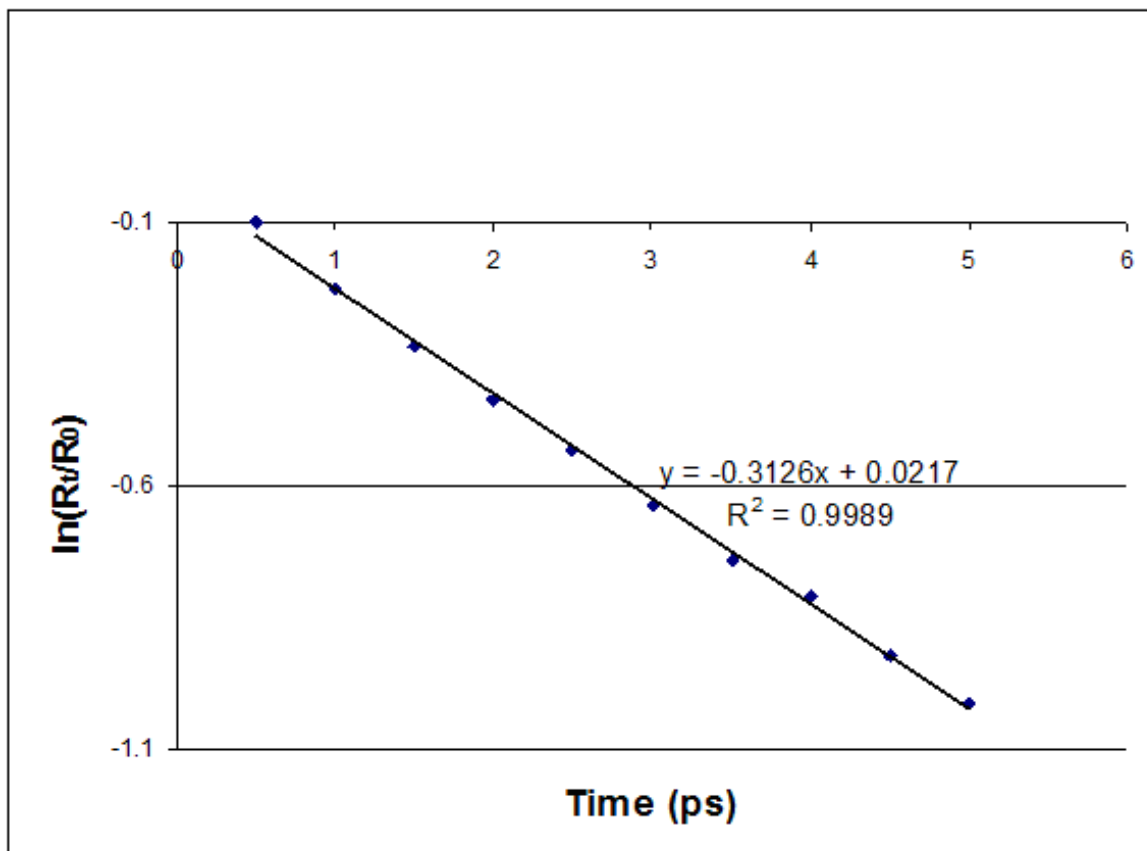


Figure III.6 First-order decay plot of *trans*→*cis* isomerization when OH mode is excited (the internal energy is 1.70 eV, including zero point energy).

The detailed rate coefficients and corresponding standard deviations are given in Table III.6. It is clearly found that the reaction rate is nearly independent of excitation types. In our result, the rate coefficients vary by a factor of 2.4 for the *cis*→*trans* isomerization from 0.210 ps⁻¹ to 0.505 ps⁻¹. Conducting similar investigations on the empirical surface, the factor was found to be as high as 22.8.³ The *trans*→*cis* isomerizations rate coefficients in our study was found to vary by a factor of 1.82 from 0.199 ps⁻¹ to 0.363 ps⁻¹. The factor of *trans*→*cis* isomerization reported by Guan and

Thompson³ is about 7.0. These results suggest that some major differences exist between MD on an empirical surface and on a NN *ab initio* surface.

The effectiveness of exciting a vibrational mode on the NN surface is different from the empirical surface employed by Guan and Thompson.³ In their study, HON excitation results in the highest reaction rate constant, while the excitation of OH stretch gave the lowest. In contrast, our study shows that the ONO bend is the most effective mode in *cis-trans* isomerizations, and HON is the least effective.

Table III.6 *Cis-trans* isomerization rate constants for the total energy of 1.70 eV
(including zero point energy)

Reaction	Mode excited	Rate coefficient (ps ⁻¹)	Standard deviation (ps ⁻¹)
<i>cis-trans</i>	OH	0.265	0.005
	N=O	0.491	0.014
	HON	0.210	0.007
	O-N	0.405	0.010
	ONO	0.505	0.011
	All	0.401	0.009
<i>trans-cis</i>	OH	0.199	0.003
	N=O	0.303	0.012
	HON	0.334	0.011
	O-N	0.363	0.013
	ONO	0.260	0.010
	All	0.255	0.010

The most important difference between dynamics on the NN and empirical surfaces is the extent of intra-mode coupling. The reactions investigated on the NN surface are observed to have much greater intra-mode coupling than those conducted on the empirical surface. As a direct consequence, it is easier for energy to spread among the

modes. From the results reported by Guan and Thompson, intra-mode coupling is hardly observed, since the reaction rate coefficients are significantly different when different modes are excited. This, in fact, is not surprising as it was mentioned that the coupling terms were omitted since they used three Morse potential terms, two bending terms, and a Fourier transformed function to describe the torsional potential.

Since the modes are highly coupled on the NN *ab initio* surface, we do not observe large differences in term of reaction rates when different modes are excited. The ratios of *cis*→*trans*/*trans*→*cis* rate coefficients vary from 0.63 to 1.94. In most cases, *cis*→*trans* rate is found to be higher than *trans*→*cis* rate when a particular mode is excited, except in the case HON bending where we observe higher rate of *trans*→*cis*. Guan and Thompson³ reported these ratios from 2.0 to 12.9.

We have also performed an investigation of *cis*→*trans* reaction on the NN committee surface when OH mode is excited. The resulting rate constant from this investigation is 0.309 ps^{-1} , which is 17% higher than the rate constant reported from conducting MD on the first NN surface, which is 0.265 ps^{-1} . Recall that the fitting error of the NN committee is 0.0111 eV, which is 35% lower than the fitting error of the first NN surface (0.170 eV).

IV. N-O bond dissociation

The potential barrier of N-O bond dissociation has been investigated using the MP4(SDQ) *ab initio* method. By extending the N-O bond from 2.2 Å to 2.8 Å, it is seen that the potential energy begins to drop at 2.78 Å. Therefore, we take this distance to be the dissociating distance. Also, it is indicated that the barrier height is 2.15 eV. As

reported by Herzberg, the zero point energies of radical HO and NO are 3735 and 1904 cm^{-1} , respectively.¹⁰² Thus, the difference in zero-point energy for $\text{HONO} \rightarrow \text{HO} + \text{NO}$ reaction is 0.18 eV. As a consequence, the minimum energy required to break N-O bond in the reaction is 1.97 eV (190 kJ mol^{-1}). It is reported by Guan and Thompson that the value on the empirical surface is 1.77 eV.³

The dissociation reaction is investigated at two different levels of internal energy, which are 3.1 eV and 3.3 eV. As previously noted, 2.78 Å is considered to be the dissociating distance; therefore, a trajectory is terminated when either 5 ps of molecular time elapse or the N-O distance reaches 2.78 Å.

Seven investigations are conducted for seven different types of excitations. In the first six investigations, excitation energy is partitioned into a vibrational mode. In the last investigation, excitation energy is partitioned equally among the six vibrational modes. Since the molecule switches from between two conformations very rapidly at these two internal energy levels, we choose to start the trajectory from *cis* configuration with the choice making no major difference.

The rate constants are obtained directly from the first-order decay plots. A typical decay plot of N-O dissociation is shown in Figure III.7. It is found that the plots are linear in all cases with excellent standard deviations. The results are shown in Table III.7 in detail. As expected, the rate constant at 3.3 eV is higher than the one at 3.1 eV when similar excitation is conducted.

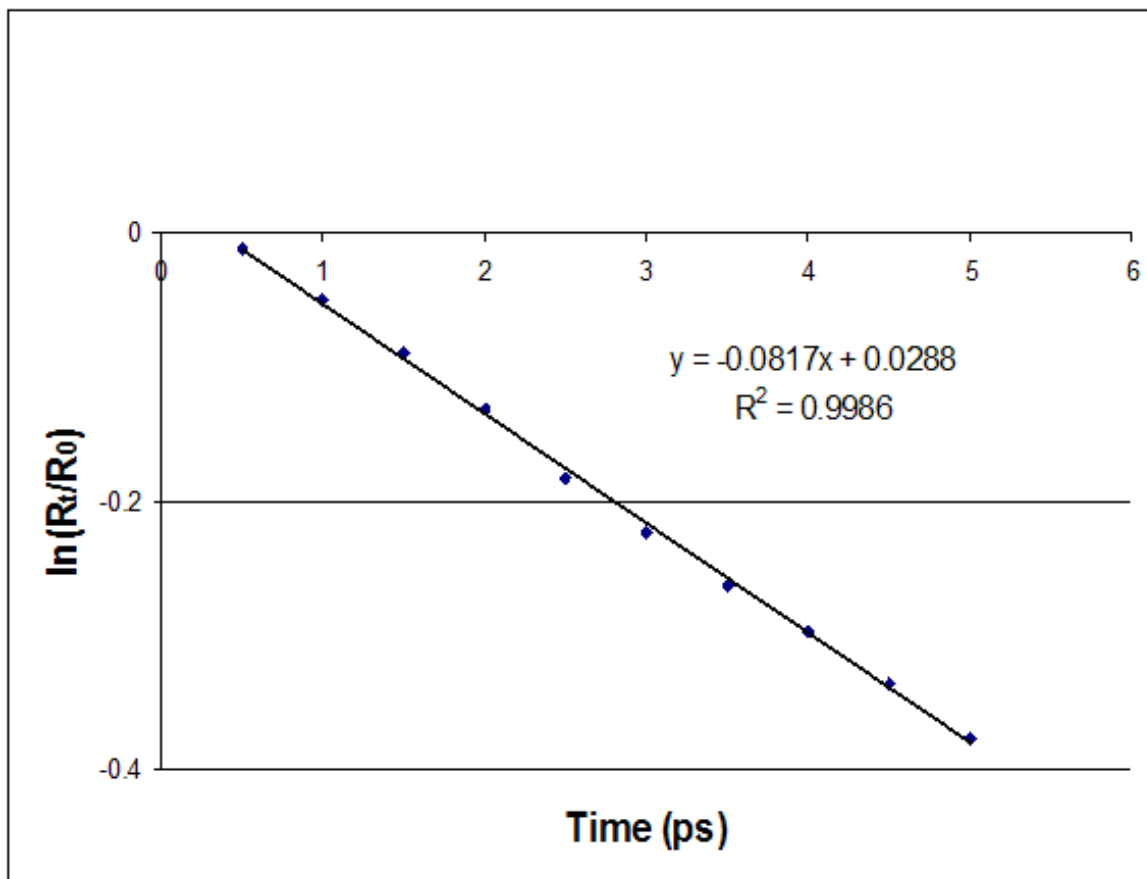


Figure III.7 First-order decay plot for N–O bond dissociation when the OH mode is excited. The total energy studied in this case is 3.10 eV (including zero point energy).

Table III.7 N-O bond dissociation calculated rate constants at total energies of 3.10 eV and 3.30 eV (including zero point energy)

Total energy	Mode excited	Rate coefficient (ps ⁻¹)	Standard deviation (ps ⁻¹)
3.10 eV	OH	0.082	0.001
	N=O	0.075	0.001
	HON	0.042	0.001
	O-N	0.165	0.003
	Torsion	0.136	0.002
	ONO	0.063	0.001
	All	0.092	0.002
3.30 eV	OH	0.151	0.002
	N=O	0.146	0.001
	HON	0.072	0.001
	O-N	0.237	0.008
	Torsion	0.212	0.002
	ONO	0.116	0.002
	All	0.157	0.004

Our results clearly indicate that the excitation of O-N mode enhances the reaction rate most effectively as expected. The excitation of torsional mode is also very effective, which results in the second highest reaction rate constants. Excitations of four other modes do not give high rate constants.

Intra-mode coupling on the NN surface is significantly greater than on the empirical surface. As reported by Guan and Thompson,³ the ratio between the highest and lowest rate constants is 69.4 at the internal energy of 2.26 eV. In our investigation, the corresponding ratio is 3.9 when the internal energy is 3.1 eV, and 3.3 when the internal energy is 3.3 eV. This, again, reflects the ability of energy to spread rapidly on the NN *ab initio* surface during the dissociation of N-O bond.

V. Summary

The *cis-trans* isomerizations and N-O bond dissociation of nitrous acid are investigated on an NN *ab initio* PES obtained by employing a feed-forward NN to fit 21,584 data points selected by novelty sampling. The electronic structure calculations are executed using MP4(SDQ) level of theory on the 6-311G(d) basis set. With five individual (6-41-1) NN fits being performed, the average absolute errors vary from 0.012 eV to 0.017 eV. When a committee of these five NNs is constructed, the average absolute error drops to 0.0111 eV as a result of nearly random error cancellation. This accuracy is better to that obtained using the IMLS methods.

Molecular dynamics trajectories are integrated on a single NN surface. The average computational time required for a trajectory is 4.8 s when calculations are performed using a 2.4 GHz processor. This computational time compares favorably to that for other methods. As the fit is done only once, gradients are calculated easily, and the parameters are not required to be adjusted at every integration step as done using the other methods. Also, the ability to fit such a large database also implies a good description of chemical intuition.

Prior to trajectory integration, a randomized configuration is generated, and excitation energy is introduced into the desired modes. Both of these steps are executed using projection methods.¹⁰⁸ This process also guarantees no center-of-mass motion and zero total angular momentum; all internal energy is distributed into the six vibrational modes of HONO.

The *cis-trans* isomerization investigations are executed at a total energy of 1.70 eV. Six different energy partitionings are performed, and all decay plots clearly indicate

the reaction to be first order (each plot is an investigation of 2,000 trajectories). As different modes are excited, the rate coefficients of *cis*→*trans* vary by a factor of 2.4 from 0.210 to 0.505 ps⁻¹, which emphasizes large intra-mode coupling on the NN surface. When MD calculations were performed on an empirical surface, the ratio was found to be 22.8 by Guan and Thompson.³ The rate coefficients of *trans*→*cis* vary by a factor of 1.82 compared to the factor of 7.0 reported by MD investigations on the empirical surface.

We have also found that the ratio of *cis*→*trans*/*tran*→*cis* rates to range from 0.63 to 1.94, depending on the mode excited. In most cases, the rate of *cis*→*trans* is always faster, except for the case of HON mode excitation. As reported previously, the *cis* conformation is a little more energetically stable than the *trans* conformation; however, it is easier to switch from *cis* to *trans* than the other way. Doing calculations on the empirical surface,³ the ratios of rates vary from 2.0 to 12.9.

When we use the NN committee, which is a combination of five individual networks, to investigate *cis*→*trans* reaction when HON mode is excited, the rate constant is 0.309 ps⁻¹. When calculations are executed on a single NN, the result is reported as 0.265 ps⁻¹. The fitting error of the NN committee is 35% less than the fitting error of the single NN, and the rate constant drops 17%. The computational time for MD when the NN committee is employed is almost five times as required for calculations on a single network, but still better than the other methods.

The N-O bond dissociation is investigated at two different energy levels, which are 3.1 eV and 3.3 eV. For each energy level, seven different partitionings of excitation energy are investigated. A decay plot is made in each case after obtaining information

from 2,000 trajectories. All dissociations are found to be first order. The ratio of maximum/minimum rates at 3.1 eV and 3.3 eV are 3.9 and 3.3, respectively.

From our results of *cis-trans* isomerizations and N-O dissociation, high intra-mode coupling between vibrational modes on the NN *ab initio* surface is observed, which makes the rate coefficients in all cases not as significantly different as reported from MD calculations on the empirical surface.³ Since the coupling terms are not included in the empirical surface, and the only coupling comes from the anharmonicity present in the Morse potentials describing the chemical bonds, their result is not surprising.

CHAPTER IV

INVESTIGATION OF $BeH + H_2 \rightarrow BeH_2 + H$

I. Development of *ab initio* potential surface

In this work, we employ the NN method again to fit the data. Although there is only one reaction channel to be considered, the construction of BeH_3 potential is more mathematically and computationally complicated than the previous one developed for HONO, which has two reaction channels.

We employ MP2/6-311G(d,p) level of accuracy to perform electronic structure calculations for BeH_3 potential, the choice that Collins and Zhang has made and reported in a study of this reaction using the Shepard interpolation method.³³ In ground electronic state, BeH state is doublet, while H_2 is singlet; therefore, the overall state of $BeH + H_2$ is doublet. A single hydrogen atom has a doublet state, and BeH_2 is singlet, thus, the overall state of $BeH_3 + H$ is also doublet. Therefore, we consider this reaction as an adiabatic process; all potentials are reported in the doublet state.

1. Novelty sampling of configurations

The effectiveness of the novelty sampling procedure has been proven to be the case for both the HONO study²⁴ and the dissociation dynamics of vinyl bromide.^{29,30} With the same numbers of atoms involved in the system, the use of novelty sampling

procedure with some modifications to enhance effectiveness should describe BeH_3 system accurately.

Novelty sampling was first introduced by Raff and his co-workers³¹ to collect geometric configurations in twelve-dimensional hyperspace. In the study of BeH_3 , this technique is employed and we have made some significant modifications to improve the effectiveness of selecting useful configuration points. “Novelty sampling” was first applied to construct the database for CH_2CHBr .^{29,30} Since this is a very large molecule in terms of both molecular dynamics study (many open reaction channels) and dimensions of hyperspace (15 dimensions for 6 atoms), in the construction of CH_2CHBr configuration hyperspace, some criteria were developed to select data that are only useful for MD trajectories. The usefulness of a new configuration is determined by calculating the scaled minimum and average distances from that configuration to all configurations of the previous data. Subsequently, some tests are designed to qualify configurations.

Selecting data from molecular dynamics trajectories using “novelty sampling” always requires a temporary PES. Prior to any data selecting processes, a temporary PES fit of the previous data has to be available. Thus, the method cannot be self-starting, and requires some techniques to obtain the first set of data. In the study of vinyl bromide, MD trajectories on an empirical PES developed by Rahaman and Raff¹⁴ were employed to produce a first set of configurations. Similarly, Le and Raff employed an empirical surface reported by Guan and Thompson³ to study HONO.²⁴ Recently, Agrawal *et al.*⁸⁰ have introduced a new technique to make the method self-starting. In their re-investigation of HONO, a first set of data was obtained from Born-Oppenheimer molecular dynamics offered by the Gaussian suite of programs (also known as direct

dynamics). The result from that study was in good agreement with the results reported by Le and Raff,²⁴ which shows the convergence of data selection is independent of the starting data set.

In this work, the 1,300 data points were generously provided by Collins and Zhang from a previous study.³³ From the initial database, we use those points to construct the first temporary NN PES. Let us denote \mathbf{v} as a six-component vector containing six scaled input values between -1 and +1 of a configuration in the database. The scaled distance between point i and point j in the database is calculated by

$$d_{ij} = d_{ji} = \sqrt{\sum_{k=1}^6 (v_i - v_j)^2} . \quad (19)$$

The i^{th} minimum and average distances are determined as follows:

$$\min_i = \min(d_{ij}) \text{ where } j = 1..n, \text{ and } n \text{ is the total number of configurations currently present in the database, and} \quad (20)$$

$$avg_i = \frac{\sum_{j=1}^n d_{ij}}{n} . \quad (21)$$

For the first set of data points, $n = 1,300$, the probability distribution plots of the minimum and average distances of the first data set are shown in Figures IV.1 and IV.2. These probability distributions are used to qualify the new configurations generated from MD trajectories in the second iteration.

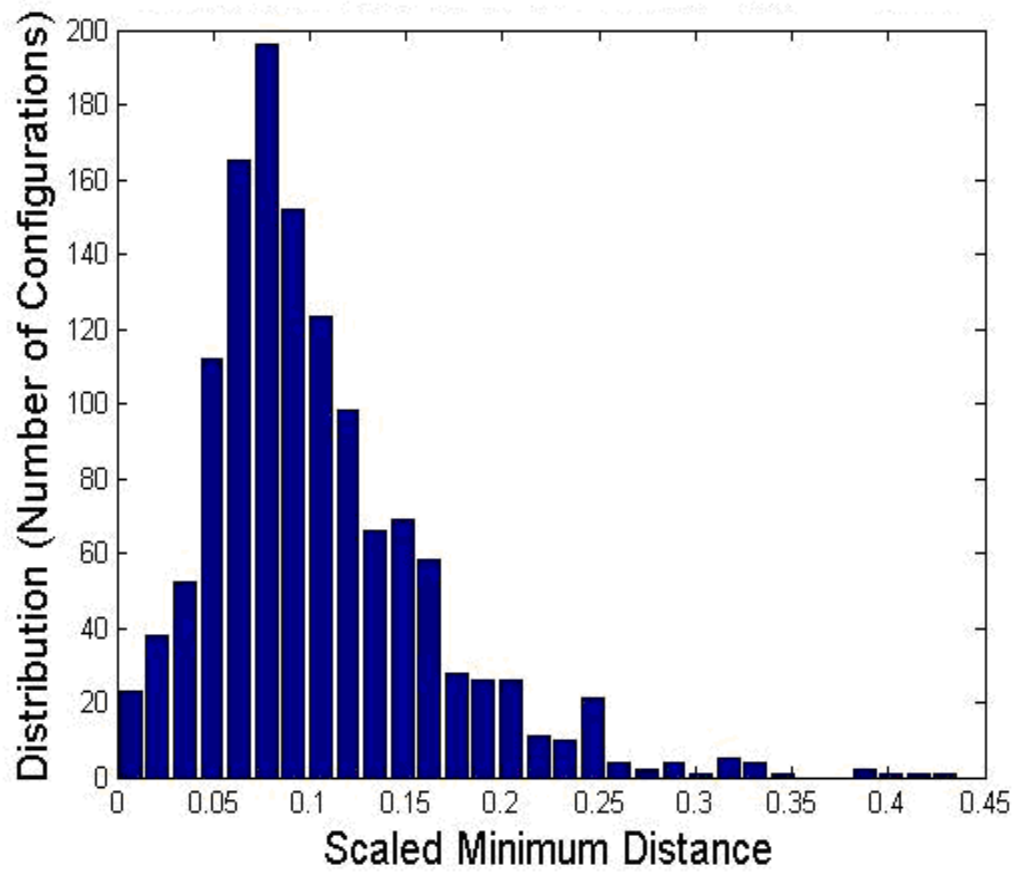


Figure IV.1 Distribution plots of scaled minimum of the first 1,300 configurations.

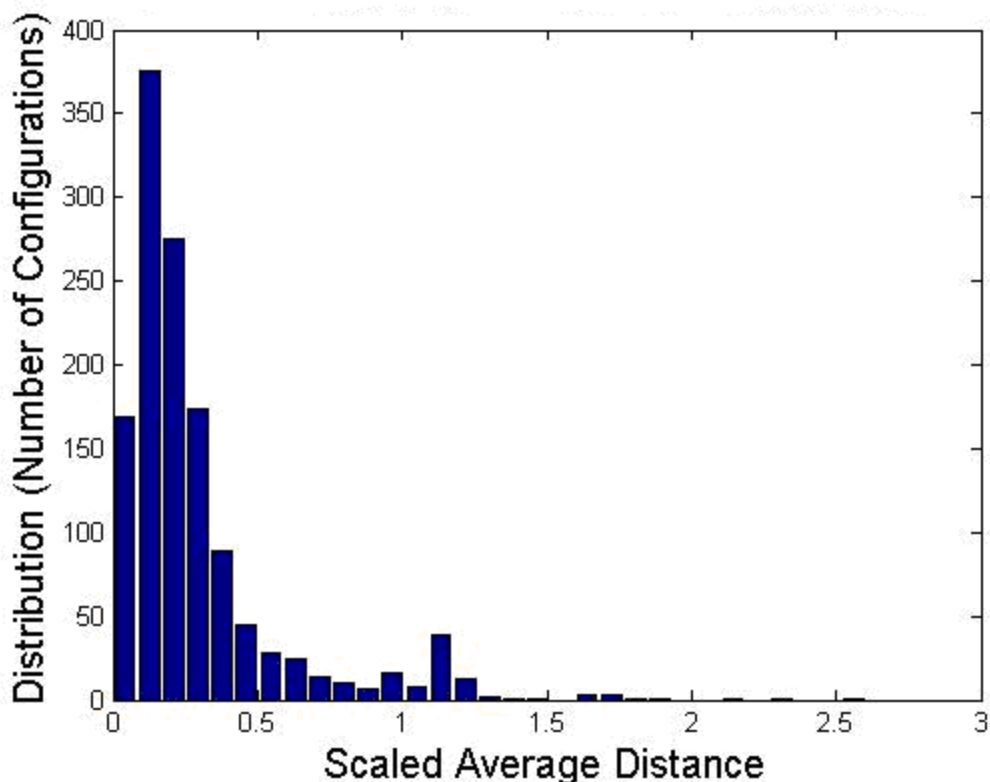


Figure IV.2 Distribution plots of average distance of the first 1,300 configurations.

A trajectory is sampled with the initial distance between BeH and H₂ centers of mass being 5.82 Å (11 a₀), and the maximum impact parameter being 0.265 Å (0.5 a₀). The technique for initializing BeH + H₂ reaction will be discussed in detail latter. As we already have the first temporary NN PES, in the first stage of selecting new data points, trajectories describing the formation of BeH₂ are executed to produce new configurations at every 0.2 time units (one time unit equals 1.018×10⁻¹⁴ s). Let us denote the minimum and average distance of a new configuration to all configurations in the previous database as min_{new} and avg_{new} . Empirically, we can state from the distribution plots of the very first 1,300 data points that there are not sufficient configurations at the minimum distance less

than 0.04 or greater than 0.13, and the average distance greater than 0.3. Therefore, a new configuration is qualified to the second stage if

- a) $min_{new} < 0.04$ or
- b) $min_{new} > 0.13$ or
- c) $avg_{new} > 0.3$

We have selected 5,181 new configurations using these criteria. Subsequently, *ab initio* calculations are executed for these configurations. A new configuration is finally added to the database if the difference between its NN energy (predicted by the temporary surface) and the *ab initio* energy is greater than the average absolute error of the temporary fit for all previous data points. During the second iteration, we have chosen 3,178 new data points.

The process is repeated iteratively until we can no longer select a significant number of data points to add to the database. During the 7th iteration, we have selected 5,027 points for *ab initio* calculations. After making comparisons between the NN-predicted energies and *ab initio* energies, 484 configurations are added to the database (only 9.6%). At this point, we determine to terminate the configuration sampling process, and the database is believed to have converged because only a few points can be added to the systems. The detailed numbers of selected data points during this sampling process are reported in Table IV.1. The database after 7 iterations contains 9,642 configurations. However, it is recognized that the points that have high potential energies would cost more effort for the NN to fit, but they are not useful for molecular dynamics studies. Therefore, 38 points are eliminated from the database, which make the number of points

become 9,604 configurations. By switching r_4 and r_5 , r_2 and r_6 due symmetric considerations, the final database contains 19,208 configurations.

Table IV.1

Detailed numbers of configurations selected during the data sampling process

Iteration	Number of points for <i>ab initio</i> calculations	Number of points added to the database	Percentage	Total points in the database
1	0	0	0	1,300
2	5,181	3,178	61.3%	4,478
3	5,846	1,853	31.7%	6,331
4	5,124	844	16.5%	7,175
5	5,195	728	14.0%	7,903
6	5,085	1,255	24.7%	9,158
7	5,027	484	9.6%	9,642

2. NN fitting of BeH_3 potential

The number of neurons in the hidden layer is selected in accordance with the complexity of the system under consideration. For a six-body molecule like vinyl bromide (CH_2CHBr),^{29,30} the number of neurons used was 140. Less complicated molecules, such as HONO,^{24,80} requires 40 neurons or more to fit the data. In this study, the configuration space is too sparse (two diatomic molecules are sampled with centers-of-mass far apart). Consequently, we employ a NN with 60 neurons in the hidden layer. Once an appropriate number of hidden neurons is chosen, the fitting process is done only once, and all parameters are saved for MD investigation.

Our data are divided into 3 sets: training, testing, and validation sets. Empirically, the training set consists of 80% of the data, while the other two sets consist of about 10% each. The maximum number of training iterations (epochs) is 1000. The *Levenberg–*

Marquardt algorithm is applied to minimize the mean square error of the training set.²² During the training process, an important issue to be concerned about is over-fitting.⁷³ To prevent over-fitting, an empirical technique known as “early-stopping” is used to terminate the training process as the mean square error of the validation set increases in 6 consecutive epochs.

The actual potential surface of BeH_3 is a committee of five feed-forward NN's. After fitting each different NN individually by selecting different sets of training data, the NN committee is calculated by taking the average of the five networks. Recently, a NN committee has been employed in the HONO study.²⁴ The reported fitting error of the NN committee is significantly less than the errors of any individual network as a result of canceling nearly random errors. In the study of BeH_3 , the average absolute error for each individual network varies from 0.0046 eV to 0.0051 eV. The average absolute error of the NN committee is 0.0046 eV, which is slightly improved in comparison to the errors of individual networks (as shown in Table IV.2). The use of NN committee in this study does not help to cancel many nearly random errors; in fact, it has shown that the training process really converges when we investigate the outputs of all neural networks, and confirm the difference between NN outputs is extremely small. The absolute error distribution of all data is shown in Figure IV.3.

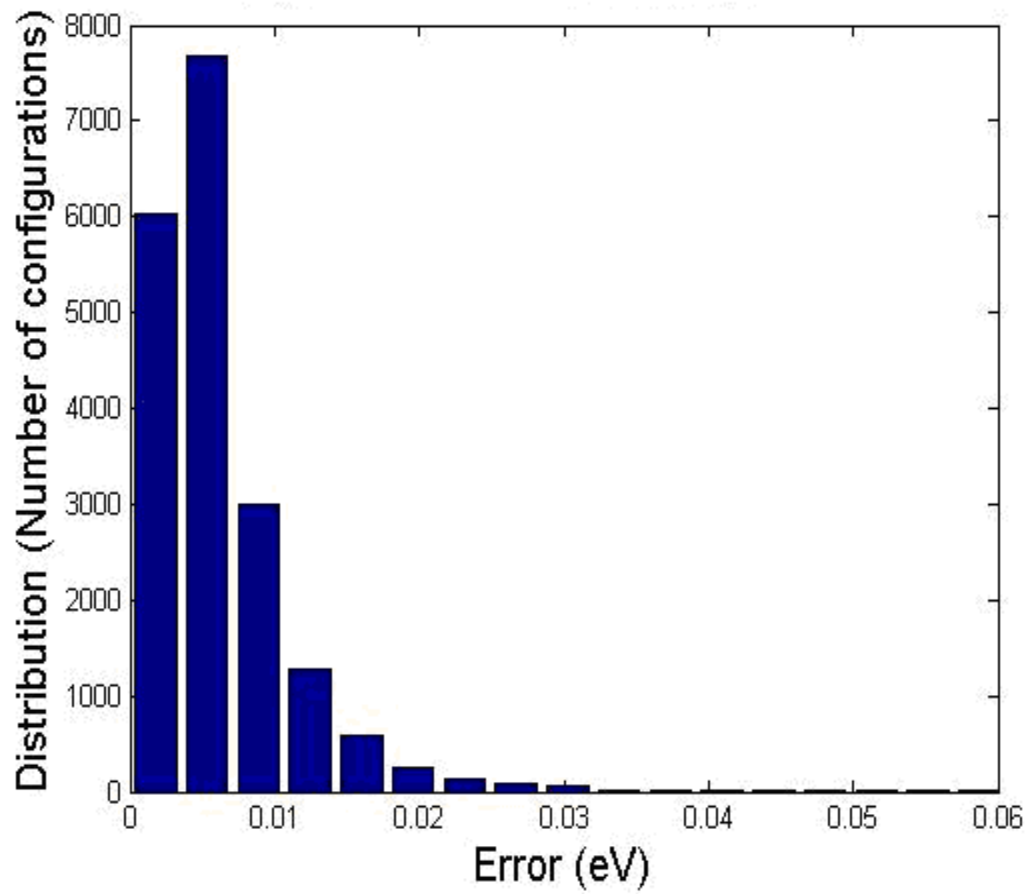


Figure IV.3 Absolute error distribution of all data of NN committee. The number of small errors (< 0.01 eV) dominate the distribution, which results in a good average absolute error.

Table IV.2

Average absolute error of the networks

Network	Average absolute error (eV)
1	0.0046
2	0.0051
3	0.0046
4	0.0049
5	0.0047
NN Committee	0.0046

The gradients calculated by the NN committee are very important in MD investigations. Although they are not directly fitted by the NN, we expect them to be well predicted. A set of 965 points is chosen as a gradient testing set. The energy gradients with respect to 12 Cartesian coordinates are computed, and compared to the true MP2 gradients. The average absolute error is $0.026 \text{ eV } \text{\AA}^{-1}$, while the average percent error is 0.17% in a range of $-7.66 \text{ eV } \text{\AA}^{-1}$ to $+7.66 \text{ eV } \text{\AA}^{-1}$. Figure IV.4 shows a histogram of gradient absolute errors. A dominating number of small absolute errors reveal that not only the outputs are fitted very well, but the gradients are also excellently predicted.

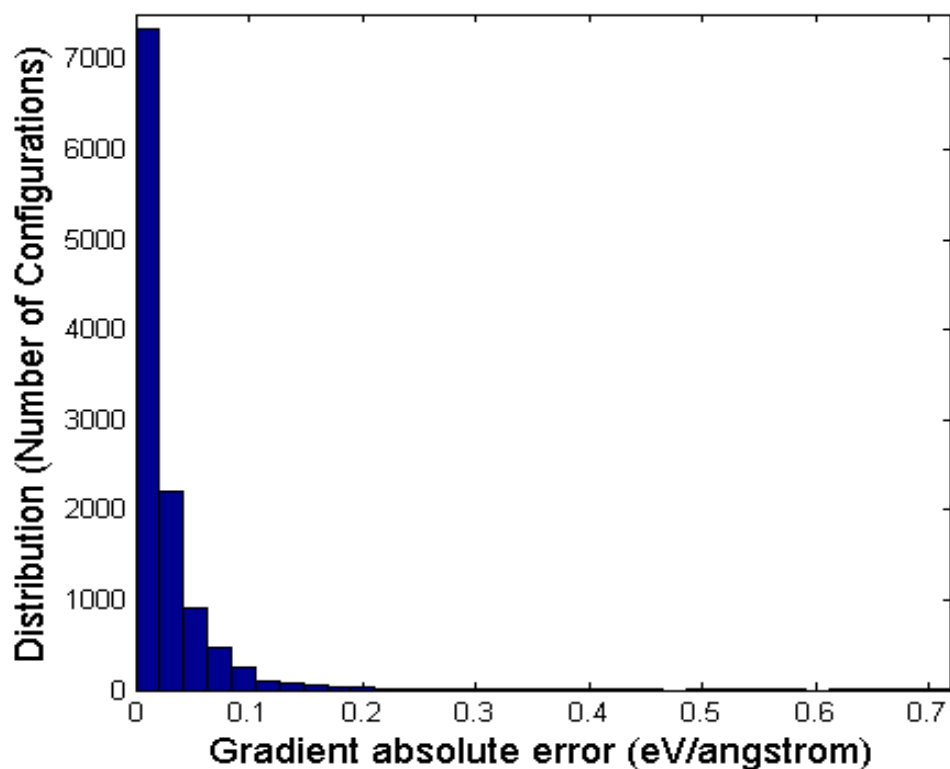


Figure IV.4 Distribution of gradient absolute errors.

II. Classical dynamics investigation of $\text{BeH} + \text{H}_2 \rightarrow \text{BeH}_2 + \text{H}$

1. Initializing the $\text{BeH} + \text{H}_2$ reaction

Before executing classical trajectories of the reaction, BeH and H_2 have to be sampled individually, assuming that there is no interaction between the two molecules at a relatively large distance. In order to do this effectively, we have constructed two simple *ab initio* PES's for BeH and H_2 .

BeH has a doublet electronic configuration in the ground state. *Ab initio* calculations are executed at MP2 level of theory using the 6-311G(d,p) basis set with various atomic distances between Be and H from 0.9 Å to 3.15 Å. A feed-forward NN fit

of BeH potential is performed with 4 neurons in the hidden layer, giving the form to the potential function as

$$V_{BeH}(r) = \left\{ \sum_{i=1}^4 w_2^{BeH} \tanh \left[w_{1i}^{BeH} \left(2 \frac{r - r_{\min}^{BeH}}{r_{\max}^{BeH} - r_{\min}^{BeH}} - 1 \right) + b_1^{BeH} \right] + b_2^{BeH} + 1 \right\} \frac{V_{\max}^{BeH}}{2} \quad (22)$$

where $r_{\max}^{BeH} = 3.15 \text{ \AA}$ and $r_{\min}^{BeH} = 0.9 \text{ \AA}$, and $V_{\max}^{BeH} = 2.35 eV$. The values of w_1^{BeH} , w_2^{BeH} , b_1^{BeH} , and b_2^{BeH} are reported in Table IV.3. We obtain an extremely small absolute average error of 0.0008 eV. The plot of the fit is illustrated in Figure IV.5.

Table IV.3 Fitting parameters of BeH potential surface

w_1^{BeH}	b_1^{BeH}	w_2^{BeH}	b_2^{BeH}
6.25	-5.64	-0.006	8.18
2.62	0.30	0.817	
4.14	2.09	0.186	
3.51	4.42	-8.584	

All parameters are unitless

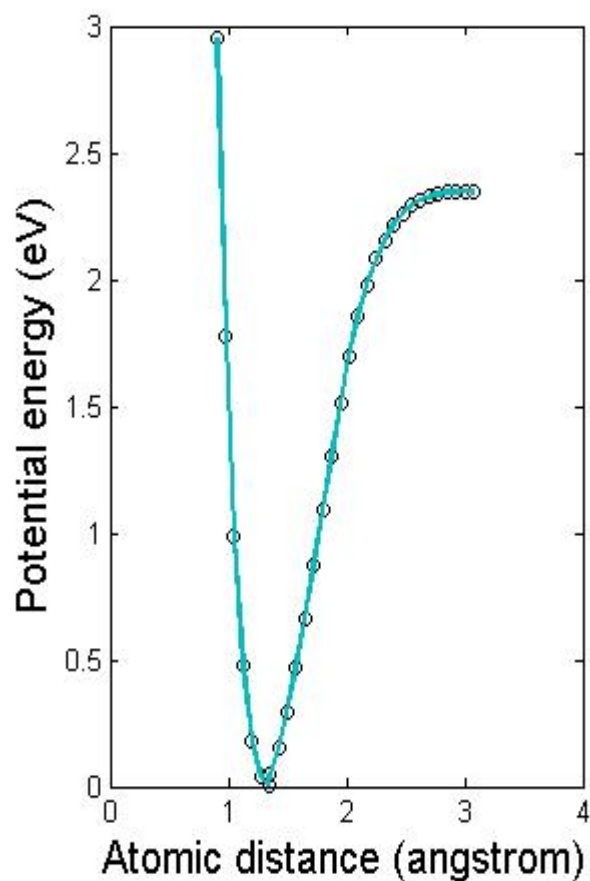


Figure IV.5 Potential energy curve of BeH. The solid curve represents the fit by formula (22), and the circles (o) represent the real *ab initio* potential energies of BeH.

BeH molecule is initially assigned at its equilibrium position (with $r = 1.34 \text{ \AA}$). A ground state vibrational energy of 0.127 eV (12.3 kJ mol^{-1}) is introduced, and the trajectory is integrated for a random period of time. Subsequently, the molecular Cartesian coordinates and momenta are rotated about the y-axis by a uniformly randomized angle from 0 to π , and then rotated about the z axis by a uniformly randomized angle from 0 to 2π .

Hydrogen gas (H₂) is in singlet state at the ground state level. We also execute *ab initio* calculations for H₂ using the same level of theory at various atomic distances from 0.45 Å to 1.54 Å. The potential of H₂ has the form

$$V_{HH}(r) = \left\{ \sum_{i=1}^4 w_2^{HH} \tanh \left[w_{1i}^{HH} \left(2 \frac{r - r_{\min}^{HH}}{r_{\max}^{HH} - r_{\min}^{HH}} - 1 \right) + b_1^{HH} \right] + b_2^{HH} + 1 \right\} \frac{V_{\max}^{HH}}{2} \quad (23)$$

where $r_{\max}^{HH} = 1.54$ Å and $r_{\min}^{HH} = 0.45$ Å, and $V_{\max}^{HH} = 3.45$ eV. The values of w_1^{HH} , w_2^{HH} , b_1^{HH} , and b_2^{HH} are shown in Table IV.4. The reported fitting error is 0.0004 eV, and Figure IV.6 illustrates the fitting potential energies on the *ab initio* energies.

Table IV.4 Fitting parameters of H₂ potential surface

w_1^{HH}	b_1^{HH}	w_2^{HH}	b_2^{HH}
-1.04	0.920	-1.181	15.33
-1.65	0.146	-0.508	
-2.63	-0.845	-0.130	
-2.40	-3.631	15.068	

All parameters are unitless

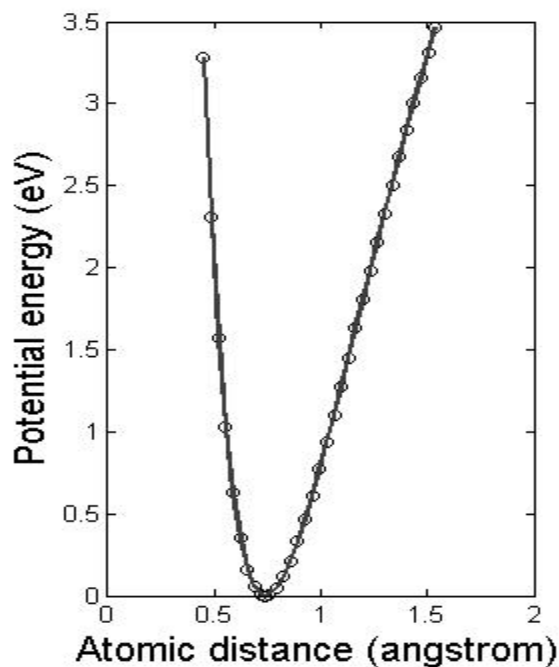


Figure IV.6 Potential energy curve of H₂. The solid curve represents the fit by formula (23), and the circles (o) represent the real *ab initio* points of H₂ potential.

The sampling process of H₂ is conducted similarly to the sampling process of BeH with a zero point energy of 0.261 eV (25.2 kJ mol⁻¹) being used for H₂ trajectories. After H₂ is sampled, the two molecules are separated by a distance of 5.82 Å (distance between the two centers of mass), and a random impact parameter with the maximum value of 0.265 Å is applied. An amount of translational energy inserted to the system with no center-of-mass motion.

2. Classical dynamics investigation

According to a previous study by Collins and Zhang,³³ the reaction probability derived from quantum dynamics calculations³⁴⁻³⁶ has a maximum value of 0.11 for

translational energy in the range of 40 kJ mol^{-1} to 77 kJ mol^{-1} when the total angular momentum of the system is zero (the impact parameter is 0). The potential barrier is investigated in a previous study by Collins and Bettens.³⁷ As MP2/6-311G(d,p) level of theory is executed along the reaction coordinate, the potential barrier is reported to be 51 kJ mol^{-1} (0.529 eV).

In the first part of this study, we conduct an investigation using classical dynamics trajectories at zero impact parameter for purposes of comparison. The translational energy is varied from 0.415 eV to 0.829 eV . The classical Hamilton's equations of motion are solved numerically using the fourth-order Runge-Kutta method,¹⁰⁹ and the total energy is conserved to the fifth digit. A trajectory is terminated when one of the two criteria given below is satisfied:

- The distance between BeH and H-H centers of mass keeps increasing and reaches 5.83 \AA . In this case, no chemical reactions are found, the two molecules just bounce off each other.
- The distance between Be and one H from the H_2 molecule is less than 1.5 \AA , and the distance between two hydrogen atoms in H_2 is greater than 2.5 \AA . In this case, the formation of BeH_2 is found.

For each translational energy level, 5,000 trajectories are studied. The reaction probability values at various energy levels are recorded (the plot of reaction probability is shown in Figure IV.7).

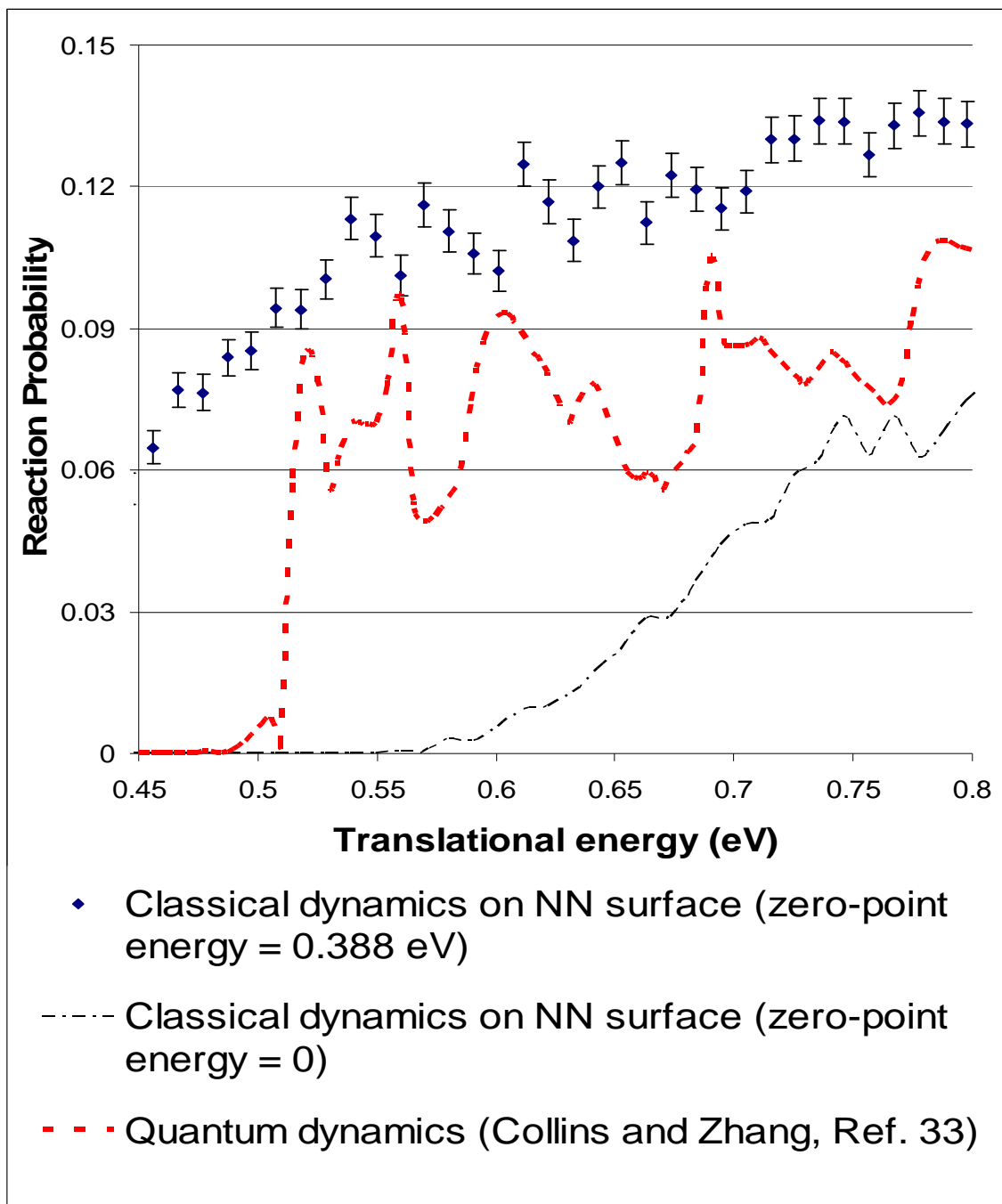


Figure IV.7 Plot of reaction probability vs. various translational energy when both gases are at their ground vibrational states ($v=0$). The reaction is investigated when BeH and H_2 both have no rotations ($j=0$), and the total angular momentum is 0 ($J=0$). The quantum dynamics results are reported by Collins and Zhang.³³

Since we conduct a classical dynamics study, there is no resonance effect shown in the reaction probability curve. In fact, the probability tends to increase as the translational energy increases. When comparisons are made, it is much more probable for the reaction to happen when classical dynamics on the NN surface is employed as a result of zero-point energy being available to promote the reaction in classical calculations. At the lowest investigated translational energy, the reaction probability is 0.043 ± 0.003 . At the highest translational energy studied, we observe the reaction probability to be 0.152 ± 0.005 , which is higher than the highest reaction probability using quantum scattering dynamics reported by Collins and Zhang (0.11).³³ The reaction threshold is also investigated by removing all zero-point energy from BeH and H₂ (this cannot be executed in a quantum mechanical system; however, since we are conducting classical dynamics study, this might be a helpful approach to search for the reaction threshold). When a translational energy of 0.56 eV is applied, we begin to observe some reactions. Therefore, 0.56 eV is concluded to be the threshold on this particular NN surface, apparently higher than the potential barrier reported by Collins and Bettens³⁷ because of very poor statistics near threshold.

The dissociation of H₂ during the trajectory pathway plays an important role in the formation of BeH₂. As a matter of fact, it is one of the two necessary conditions for reaction. This is seen clearly when H₂ is promoted to its first excited state. A second investigation is conducted with a maximum impact of 0.265 Å. BeH in this case is still assigned at its ground state, while H₂ is in the ground state or the first vibrational excited state with an internal energy of 0.783 eV (recall that the ground state vibrational energy of H₂ is 0.261 eV). It is observed that the reaction cross section when H₂ is excited is

much greater when H₂ is in its ground vibrational state (Figure IV.8). Moreover, it is observed that the probability seems to increase linearly with the applied translational energy. At the highest translational energy being studied, more than 47% of the samples are converted to the product side (the result is also shown in Table IV.5). When H₂ is at the ground state, the cross sections vary from 0.007 Å² to 0.030 Å². On the other hand, when H₂ is promoted to the first excited state, the cross sections vary from 0.05 Å² to 0.10 Å². When H₂ is in its excited state, the cross section tends to increase linearly with the translational energy up to the highest energy investigated (0.829 eV).

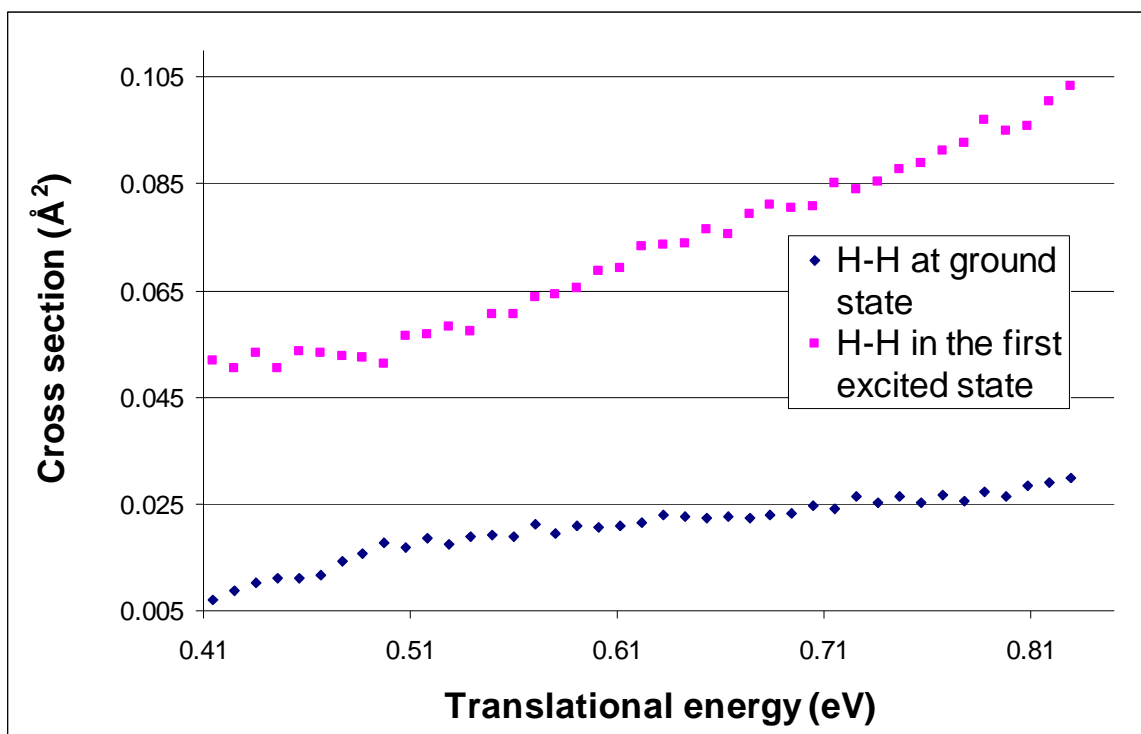


Figure IV.8 Plot of reaction cross section vs. various translational energy when H₂ is at the ground state or first excited state. The reaction is investigated when BeH and H₂ are both at their ground rotational state ($j=0$). The impact in this investigation is 0.265 Å.

Table IV.5 Reaction probability and cross section of $BeH + H_2 \rightarrow BeH_2 + H$. The vibrational and rotational quantum states of BeH are zero in all cases. H_2 gas is either at the ground state or the first vibrational excited state.

Translational energy (eV)	Reaction Probability		Cross section	
	$v = 0, j = 0$	$v = 1, j = 0$	$v = 0, j = 0$	$v = 1, j = 0$
0.41	0.032	0.007	0.236	0.052
0.44	0.047	0.010	0.242	0.053
0.45	0.051	0.011	0.229	0.050
0.46	0.050	0.011	0.243	0.053
0.47	0.053	0.012	0.242	0.053
0.48	0.065	0.014	0.239	0.053
0.49	0.072	0.016	0.239	0.053
0.50	0.080	0.018	0.233	0.051
0.51	0.077	0.017	0.256	0.056
0.52	0.084	0.018	0.259	0.057
0.53	0.079	0.017	0.264	0.058
0.54	0.086	0.019	0.261	0.057
0.55	0.087	0.019	0.276	0.061
0.56	0.086	0.019	0.275	0.060
0.57	0.097	0.021	0.290	0.064
0.58	0.088	0.019	0.293	0.064
0.59	0.095	0.021	0.298	0.066
0.60	0.093	0.020	0.312	0.069
0.61	0.095	0.021	0.315	0.069
0.63	0.104	0.023	0.334	0.074
0.64	0.103	0.023	0.336	0.074
0.65	0.102	0.022	0.348	0.077
0.66	0.103	0.023	0.343	0.075
0.67	0.102	0.022	0.361	0.079
0.68	0.104	0.023	0.368	0.081
0.69	0.106	0.023	0.366	0.081
0.70	0.112	0.025	0.368	0.081
0.72	0.109	0.024	0.387	0.085
0.73	0.120	0.026	0.381	0.084
0.74	0.115	0.025	0.389	0.085
0.75	0.120	0.026	0.399	0.088
0.77	0.121	0.027	0.414	0.091
0.78	0.116	0.026	0.421	0.093
0.79	0.125	0.027	0.441	0.097
0.80	0.120	0.026	0.432	0.095
0.82	0.131	0.029	0.456	0.100
0.83	0.136	0.030	0.470	0.103

III. Summary

An *ab initio* PES using NN fitting that describes the reaction $BeH + H - H \rightarrow BeH_2 + H$ is developed in this study. We have successfully constructed an *ab initio* potential surface for a four-body system that describes a collision of two molecules leading to the formation of products. The electronic structure calculations are executed at MP2/6-311G(d,p) level of accuracy. Previously, several potential surfaces have been developed for molecular dissociations, and in those cases, the atomic distance ranges are more restricted. In the BeH_3 system, the maximum extension of atomic distances is up to 8.1 Å; therefore, it is more difficult for the NN to generalize the fit, although the BeH_3 does not have a complicated electronic configuration.

The NN in this study requires 60 neurons in order to sufficiently fit the data. A NN committee is constructed by taking the average of five individual networks. Each network is trained by choosing the training set randomly in the database (approximately 80% of the data). The average absolute error for each individual network ranges from 0.0046 eV to 0.0051 eV (0.44 kJ mol⁻¹ to 0.49 kJ mol⁻¹), while we obtain an average absolute error of 0.0046 eV (0.44 kJ mol⁻¹) for the NN committee. This fitting error is very comparable to the testing error reported by Collins and Zhang,³³ which is 0.41 kJ mol⁻¹. Recall that our database is almost 15 times bigger, which implies the NN method can fit large molecular system that has many data points without costing too much computational time on MD calculations (the average computational time is 1.07 seconds for a trajectory when we employ a CPU of 2.4 GHz clock speed).

A modified scheme of “novelty sampling” is used in this study to select data points based on MD on a temporary NN surface. After 7 iterations, we have terminated the sampling process and chosen 9,642 configurations from MD trajectories; however, 38 points of high potential energies are eliminated as they have high energies that are unimportant in the molecular dynamics calculations. This set is then duplicated to make a final database of 19,208 configurations because of symmetric consideration of BeH₃. The final database is about fifteen times larger than the database reported by Collins and Zhang.³³

The investigation of the reaction is executed using classical trajectories. With a maximum impact parameter of 0.265 Å, the reaction is investigated by executing 5,000 trajectories at a given translational energy. Translational energy levels in the range of 0.415 eV to 0.829 eV are investigated. The computed reaction cross sections vary from 0.007 Å² to 0.030 Å² when H₂ is at ground vibrational state, and 0.05 Å² to 0.10 Å² when H₂ is in the first excited vibrational state. The reaction cross section tends to increase as the translational energy increases because classical dynamics is employed.

When we set the impact parameter to 0 (no total angular momentum), the reaction probability is investigated and compared to the results reported by Collins and Zhang.³³ They have conducted the similar reaction on an analytic PES fitted by Shepard interpolation method using quantum dynamics scattering.³⁴⁻³⁶ With the same translational energy range as we use in this study, resonance effects are observed in their results. The highest reaction probability was reported to be 0.11. This is less than our maximum reaction probability of 0.152 ± 0.005 . This difference is due to the contribution of zero-

point energy being available to promote classical reaction, which is usually not found in quantum calculations due to the quantum restrictions.

The reaction threshold is found by investigating classical dynamics of $\text{BeH} + \text{H}_2$ with no zero-point energy and various translational energy levels. When the translational energy is 0.56 eV, we begin to observe the formation of BeH_2 as a product. Therefore, 0.56 eV is concluded to be the reaction threshold within the statistical accuracy of classical trajectories.

CHAPTER V

MOLECULAR DISSOCIATION OF HYDROGEN PEROXIDE (HOOH)

I. Introduction of a new sampling method

Minimizing the sum of training squared errors is one of the NN fitting principles. During the fitting process, every point in the training set has an equal contribution. Experimentally, it is noticed that the density distribution of data has a significant affect on the fitting quality. In the training process, each point in the training set has an equal role. If a particular region of the database is more densely populated than others, the fitting error in that region is minimized; however, the fitting errors in other less dense regions will be sacrificed and become large, which increases the testing error. Therefore, it is useful to develop a new sampling technique that distributes configuration points more uniformly in order to enhance fitting accuracy. It is also helpful to develop a method that makes a system self-starting, which means construction of PES of any system is not dependent on the availability of empirical and semi-empirical potential surfaces in the literature.

Since the employment of empirical or semi-empirical surfaces is not required, collecting data points using MD trajectories is not an obligation. In fact, when using MD as described by the “novelty sampling” technique,³¹ one has to confront the non-uniform distribution of data points. This is somehow handled in NS sampling as the minimum and

average distance criteria tend to prevent density from building up in one region. In low-potential-energy regions, the kinetic energy is high, and the force field acting on every atom is high; therefore, atoms would move fast. On the other hand, atoms tend to move slowly in the high-potential-energy regions, and selecting data points would result in high density distribution unless the sampling frequency is adjusted. Thus, it is necessary to make the sampling frequency compatible to the instantaneous acceleration of atoms. To deal with this issue, Pukrittayakamee *et al.*^{86,87} has proposed a method in which they adjusted the integration step based on the level of kinetics energy.

To date, most fitting methods mainly focus on the potential fits, although the potential is not required in MD investigations. The gradients of the potential with respect to input coordinates (force fields), on the other hand, are more important than the output potential itself as they are used to integrate the classical Hamilton's equations of motion. Therefore, it is more important that the force fields are accurate. In most potential surface studies, an assumption is made, namely, if the potential is well fitted, the shape of the function is mathematically well-described, and the force fields are also accurate. However, this is not necessarily true, especially when over-fitting occurs. Adopting the idea of force field fitting, Pukrittayakamee *et al.*^{87,88} have proposed an algorithm that simultaneously fits both output and gradients. Since this feature is not available in MATLAB NN toolbox, the training algorithm has to be constructed.

The valence internal coordinates describing the HOOH system consist of three bond distances (H-O, O-O, and O-H), two bending angles (H-O-O and O-O-H), and one dihedral angle. This set of coordinates is used as the input vector for NN training. The *ab initio* potential energy is used as NN target. In previous NN studies,^{24,32} all atomic

distances were used as NN inputs. In this study, we have found that the fitting error when using internal valence coordinates is lower than the error resulting from using atomic distances as NN inputs.

If the molecule under study has three atoms or less, a grid search over the entire configuration space can be performed. If the molecule contains four or more atoms, the configuration space is very large due to the number of coordinates. In such cases, more efficient sampling methods have to be developed. The well-known “novelty sampling” method is among those.

In some previous works,^{24,29,30,80,81} MD trajectories have been used to sample configurations. This idea was first adopted by Collins and Zhang.³³ The sampling process of configuration space using this method consists of two steps. First, an initial set of data is generated in order to provide a temporary NN fit. In the second step, more geometric configurations are obtained and incorporated with the previous data until the system is believed to converge.

Traditionally, to obtain the first set of data, MD trajectories on an empirical surface or chemical intuition can be employed. This method has been proven to cover the configuration space well; however, in order to execute this, an empirical surface from the literature must be available. Another method has been proposed recently by Agrawal *et al.*,⁸⁰ which makes the method self-starting. In that method, they employ direct dynamics offered by Gaussian suite of programs.

We develop the HOOH database based on selecting configurations around data points that are not well-fitted by the NN. By adding more data points around those badly-

fitted configurations, we expect the NN to focus more on that region, and give a better fit. This method is independent of MD trajectories.

In the first step, we construct a set of configurations which represents most of stationary points of the HOOH system. This is done by performing a procedure provided by Gaussian suite of program known as “partial optimization.” In a partial optimization procedure, one or more valence internal coordinates are held constants, while the remaining parameters are optimized to minimize the potential energy. The Newton-Raphson method is used to analyze the first and second derivatives matrices with respect to the internal valence coordinates iteratively until the convergence criterion is satisfied (the potential energy is numerically minimized). Beside the Newton-Raphson method, the “steepest descent” method can also be requested when using the program.

Three chemical bond distances H-O, O-O, and O-H are denoted as r_1 , r_2 , and r_3 , respectively; θ_1 and θ_2 are denoted as two bending angles, while ϕ is denoted as the dihedral angle. In each partial optimization, ϕ and one of the other five parameters are held constant in energetically reasonable ranges. Subsequently, four remaining parameters are varied to minimize the potential energy. This process is executed when ϕ is varied in a range of 0° and 180° , while r_1 , r_2 , and θ_1 are varied in those ranges given in Table V.1.

Table V.1

Number of configurations obtained from each variation. In each sampling, the dihedral angle is varied from zero to 180°, as described in the text.

Varying parameter	Number of configurations	Range
O-H bond (r_1)	722	0.71 Å – 1.91 Å
O-O bond (r_2)	704	1.12 Å – 2.92 Å
O-O-H angle (θ_1)	722	40.0° - 169.8°
Total = N_0	2148	

At the end of this process, we obtain $N_0 = 2,148$ configurations with a maximum potential energy of 4.36 eV. The detailed number of data points obtained from each variation is listed in Table V.1. In this process, the data points obtained represent most of stationary points of HOOH energetic geometry.

In the next stage, more data points are collected to cover the hyperspace that is energetically useful for the chemical dissociation of HOOH. We begin this process by scaling all the inputs and output in the range of -1 to +1 as previously described, and employ these scaled coordinates to sample more configurations.

In the regions of six dimensional hyperspace where there are insufficient data points to describe the surface accurately, it is very likely that the NN extrapolates and results in bad fitting. Consequently, the NN-predicted first derivatives with respect to the input coordinates (bond distances, bending angles, and a dihedral angle) will exhibit significant differences compared to the derivatives calculated by the *ab initio* method (MP2) even though the outputs in that region are well-predicted. Adopting this idea, we

are able to identify the regions lacking a sufficient number of data points. Prior to searching for data points, a temporary NN fit is performed. Then, all gradients with respect to input coordinates are calculated using the fitted function. The NN-predicted gradients are compared to the corresponding MP2 gradients. If the percent difference is higher than 3%, we perform a configuration search around that point as follows:

- A unit vector v is randomly generated. A new point p_{new} is computed as $p_{new} = p + k \cdot grid \cdot v$, where $grid$ is the average distance between the data in the first set. For the present HOOH system where the initial database contains 2,148 points, $grid = 0.119$ in scaled units. k is a scalar random number between 2 and 3.
- If the distance between p_{new} and any other point in the data base is smaller than $grid$, the point p_{new} is discarded. In addition, the NN potential at point p_{new} is computed using the temporary NN fit. If the computed potential is greater than 4.36 eV, p_{new} is rejected; otherwise, it is saved for *ab initio* calculations and a final qualification test.
- Since the HOOH configuration hyperspace is six dimensional, this process is executed for 2^6 times around point p . Once a new point p_{new} is selected, the searching for configurations around p is terminated.

In the above procedure, a *grid* is applied to make the database more uniform by preventing new configurations being too close to the present points in the current database. The second requirement that the energy of a point has to be lower than a

maximum value is used to limit the volume of HOOH hyperspace. This condition makes the sampling process more feasible.

With $N_0 = 2148$ points, we begin to search over the database to find points that fail the derivative analysis. The process of finding new points is then executed, and N_1 points are generated. We continue to search for more points around these recently-found N_1 points by using the same procedure, and find N_2 points, then more configurations are generated from N_2 . This process is continued until no more points can be generated, or the total of $(N_1+N_2+N_3\dots)$ exceeds 5,000.

Subsequently, *ab initio* calculations are performed on the generated configurations at MP2 level. The computed energy of a new point is then compared to the NN-predicted energy. If the percent difference is less than 1%, the point is discarded as it is already well-described by the NN. Otherwise, it is included.

The convergence of the sampling process is validated by testing (1) the six fundamental vibrational wave numbers predicted by the NN surface against the MP2-predicted wave numbers and (2) the potential energy error of several hundreds of points generated from a MD trajectory. Also, the convergence can be recognized by the computational time used to search for data points (if the system is closely converged, more configurations can hardly be found in the defined region). At the termination of our sampling process, the NN surface exhibits very good agreement between MP2 and NN-calculated wave numbers. Table V.2 gives some sets of HOOH fundamental wave numbers resulted from various calculations.

Table V.2 Fundamental vibrational wave numbers resulting from various calculations

Mode description	Wave number (cm ⁻¹)			
	NN predicted	MP2/ 6-31G*	B3LYP/ cc-pVTZ	PCPSDE ¹⁵
Symmetric OH stretching	3723	3740	3766	3778
Antisymmetric OH stretching	3714	3738	3765	3762
Symmetric OOH bending	1456	1463	1438	1453
Antisymmetric OOH bending	1284	1323	1321	1297
O-O stretching	929	928	954	889
Torsion	331	338	366	392

Four iterations have been executed to construct a final database of 12,804 points. In the first iteration, 4,314 configurations are identified for MP2 calculations; however, only 1,803 points are selected. During four iterations, we have identified 16,270 points, and 10,656 are added to the final database, which means we have rejected about 34.5% to make the database more uniform. The numbers of identified, accepted, and rejected points are shown in Table V.3.

Table V.3 Number of new configurations identified by sampling process described in the text, total number rejected and accepted due to energy differences between MP2 and NN calculations in the selection process.

Iteration #	# Points Identified	# Points Accepted	# Points Rejected
1	4314	1803	2511
2	5001	3607	1394
3	4997	3738	1259
4	1958	1508	450

Totals	16270	10656	5614

When combined with 2,148 configurations from the initial set, the database contains 12,804 points after four iterations. This set is then duplicated by switching r_1 and r_3 , θ_1 and θ_2 to produce a final set of 25,608 configurations due to the symmetry of HOOH molecule.

The present sampling method allows configurations to be added by comparing the NN-predicted and MP2 gradients. As an advantage, this essentially makes the NN fit both the energy output and its gradients, thereby producing a better fit. The average absolute error of 0.0060 eV (0.58 kJ mol⁻¹) reported in this study is among the best fitting errors for a four-body system in the literature. We also find excellent agreement between fundamental wave numbers from MP2 calculations and the normal mode analysis of the NN function.

II. Electronic structure calculations

The *ab initio* calculations in this study are performed at MP2/6-31G* level of accuracy using the Gaussian suite of program.⁹³ For dissociation into two OH radicals, this level of *ab initio* calculations is not expected to give good accuracy. This PES is obtained in order to give a realistic test of the sampling method; therefore, we choose a small basis set with appropriate level of accuracy (MP2) to perform this illustrative example with small computational time.

As reported latter, the potential barrier of O-O dissociation is investigated using several different *ab initio* methods. These methods (beside MP2) are HF and MP4(SDQ) on the 6-31G* basis set, and B3LYP/cc-pVTZ.

III. Fitting methods

We have successfully trained a committee of five NNs for the database of 25,608 data points. Each NN is a feed-forward NN with two layers. The number of hidden neurons used in this study is 34. Recall that for a four-body system like HONO with two open reaction channels (*cis-trans* isomerization and N-O bond dissociation),²⁴ it requires 41 neurons to perform the fit. BeH₃ only has one reaction channel; however, the hyperspace volume of this system is much larger because of the distance between the two centers of mass. As a result, we have used 60 neurons to fit the BeH₃ system with a comparable fitting error to that reported by Collins and Zhang using the Shepard interpolation method.³³

Five individual networks are fitted using 80% of database as training set, 10% as the testing set, and the remaining 10% as validation set. In each fit, the data are randomly

distributed into each set. The training process consists of a maximum of 500 iterations (epochs) using the Levenberg-Marquardt training algorithm.²² To prevent over-fitting, we employ “early stopping”,²² which allows the training process to be terminated if the mean square error of the validation set keeps increasing in six consecutive epochs.

Five individual NNs are trained, and it is found that the average absolute errors vary from 0.0077 eV to 0.0090 eV, while the root mean squared errors vary from 0.0118 eV to 0.0160 eV with a maximum potential energy of 4.36 eV. The fitting errors range from 0.177% to 0.211% of the energy range. The detailed fitting errors of five individual networks and the NN committee are shown in Table V.4.

Table V.4 Average absolute and root mean squared testing set errors (eV) of the five NN fits forming the NN committee.

	Fit 1	Fit 2	Fit 3	Fit 4	Fit 5	Committee
Average absolute error	0.0077	0.0107	0.0079	0.009	0.0078	0.0060
Root mean squared error	0.0118	0.016	0.0124	0.0138	0.0118	0.0099

A NN committee is produced by taking the average of the five individual NN outputs. Gradients are also calculated by taking the average of five NN gradients. The error reported on the NN committee is 0.0060 eV, which is lower than the errors of any individual NNs as a result of cancellation of nearly random errors. The training outputs are plotted against the corresponding targets in Figure V.1. In Figure V.2, the distribution of testing set absolute errors is shown.

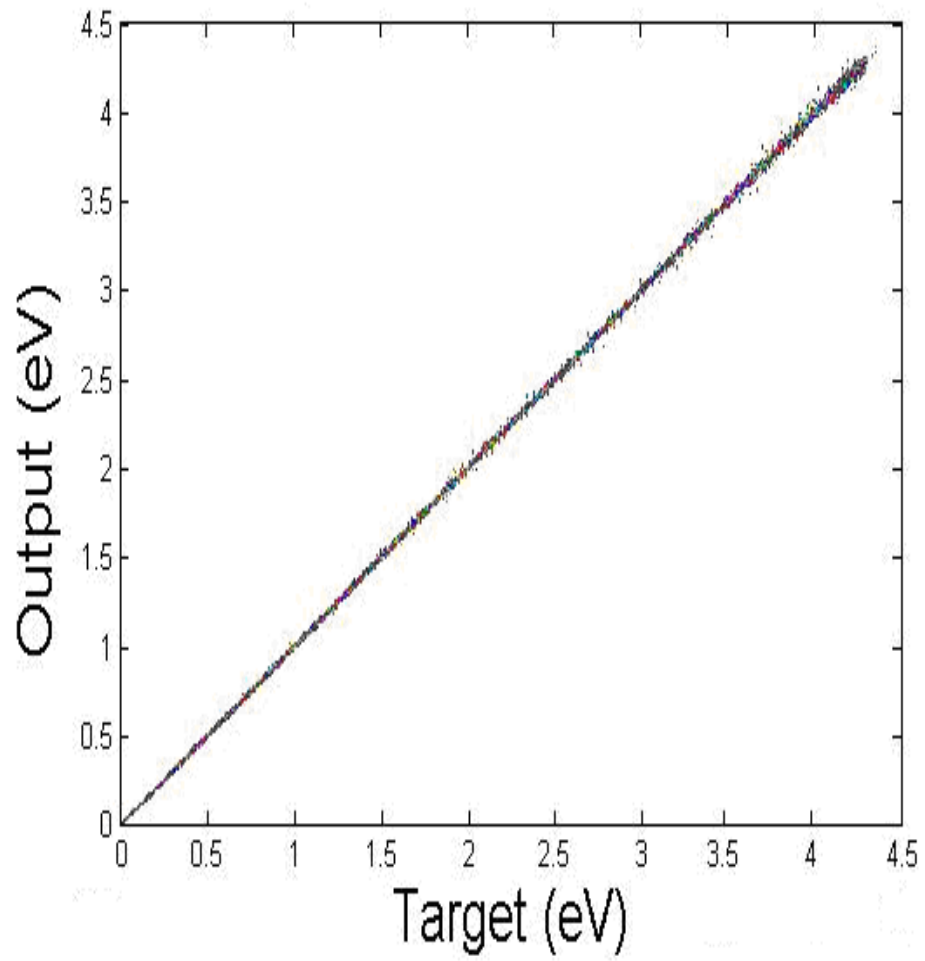


Figure V.1 Plot of targets vs. training outputs for all data points

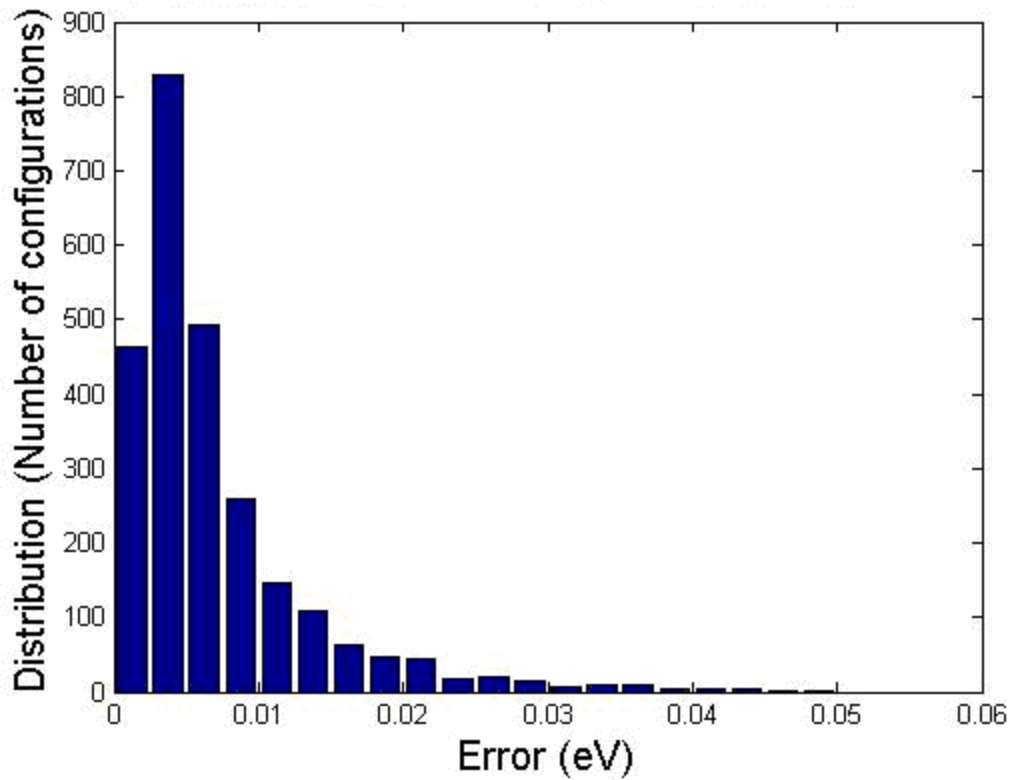


Figure V.2 Distribution of the absolute testing set errors for the NN committee. The small fitting errors (from 0 to 0.01 eV) clearly dominate the distribution. The average absolute testing set error is 0.0060 eV.

A different technique is also employed to fit the data, which is the SVM technique. As mentioned earlier, when using the support vector regression technique, there are several important parameters that have to be estimated empirically by users to utilize the best fitting result.

The most important parameter in the Gaussian exponent is γ , which determines the flexibility of the entire function. If γ is too high, each Gaussian term will be small; thereby the number of terms (support vectors) will increase as the program tries to cover the entire configuration space. The determination of γ not only decides the fitting

accuracy, but also determines the MD computational time. The “cost” value is another important parameter in the training process. The “cost” represents the maximum absolute value of α , a coefficient that exists before each Gaussian term to determine its contribution (α can be considered as a weight of each term). Higher values of c reduce the training speed because α is adjusted in a wider range. Figure V.3 shows the number of support vectors for some SVM fits with different values of c (10, 13, and 17), with various values of γ in the range of 3 to 9.

As seen from Figure V.3, the number of support vectors decreases significantly as the cost increases in a certain region. As c becomes large, the number of support vectors does not decrease significantly, but the training speed now is too slow. Therefore, the choice of very high value of c is not preferred.

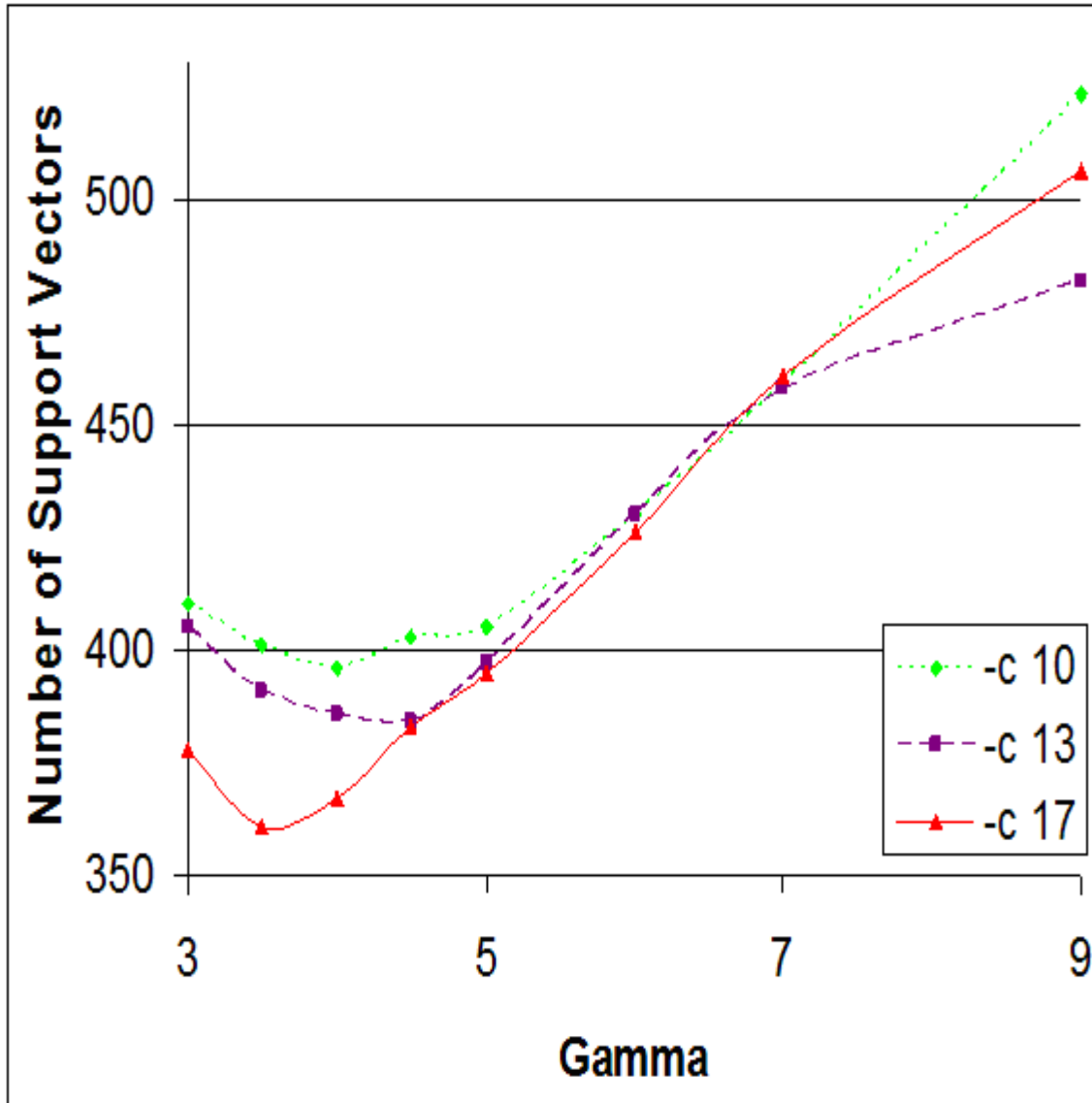


Figure V.3 The variation of the required number of SVM support vectors as a function of gamma for different cost values. When γ lies in the range from 3 to 9, the number of support vector is seen to decrease with increasing γ and the cost. The best fitting result is obtained for $\gamma = 4.0$, $c = 30$, and $p = 0.03$.

The last important parameter in an SVM training process is ρ , which is defined as the epsilon value of the “loss function.”⁹⁰ The change of this parameter has an immediate affect on the training errors. When ρ decreases, the rms error decreases; however, the number of support vectors increases, which may lead to over-fitting. To obtain the best fit (good fitting accuracy and an appropriate number of support vectors), we perform many different fits with γ varying from 3.0 to 4.5, ρ from 0.01 to 0.034. For each fit, 80% of data are randomly selected, and serve as the training set. The remaining configurations are used in the testing set. The performance of the SVM is analyzed by calculating the rms errors and performing normal mode analysis on the SVM function. Finally, we choose $\gamma = 4.0$, $\rho = 0.034$, and $c = 30$ as they give the best performance in term of reducing the number of support vectors and contributing good fitting accuracy. The number of support vectors is 447, which results in the testing rms error of 0.0384 eV, and testing average absolute error of 0.0311 eV, which is about five times higher than the two corresponding errors given by the NN committee as reported earlier.

We also present an evaluation of SVM method versus NN method, and comparisons are made in terms of MD studies. Using 34 neurons in the hidden layer, a single NN has 273 parameters totally, and the best average absolute error is reported as 0.0079 eV. On the other hand, the SVM fit requires 3130 parameters as 447 support vectors are employed. The average absolute error reported above is 0.0311 eV, which is almost five times higher than the error of a single NN. In term of trajectory time, a trajectory running for 49,113 integration steps requires about 9 s to accomplish on the NN surface, while about 2 minutes is required for the same trajectory performed on the SVM surface with less accuracy.

A further comparison is made by testing the fitting quality of a random trajectory. A random trajectory is generated, and runs for 388 steps on the NN PES with a fixed step size of $1.018 \times 10^{-16} s$ using the fourth-order Runge-Kutta integration method.¹⁰⁶ The geometric configuration and potential energy at each step are recorded during the trajectory. The SVM fit is now used to predict the potential energy of those stored configurations. The true MP2 *ab initio* energies are also calculated. The comparisons of true energies, NN energies, and SVM energies are shown in Figure V.4. The average absolute errors of the NN and SVM fit are 0.022 eV and 0.093 eV, respectively. It is shown clearly that the NN method gives better fitting generalization than the SVM method when a random trajectory test is performed despite the smaller number of parameters being employed by NN.

We have computed fundamental wave numbers by the NN surface, and compare to other *ab initio* values given by MP2/6-31G* and B3LYP/cc-pVTZ levels. Kuhn *et al.*¹⁵ used the semi-empirical PCPSDE surface to extract the wave numbers, and their results are found to be comparable to ours. All results are shown in Table V.2. The wave numbers resulted from the NN surface are in excellent agreement with the MP2 wave numbers. The wave numbers resulted from the SVM fit are not even close to the MP2 values, and are not reported here. The SVM fit has no further use in this study.

IV. Molecular dissociation of HOOH: A proof of internal hydrogen bonding

The O-O dissociation is investigated in this study using classical dynamics. Prior to executing molecular dynamics calculations, we have investigated the dissociation barrier using various *ab initio* methods, which include HF/6-31G*, MP2/6-31G*,

MP4(SDQ)/6-31G*, and B3LYP/cc-pVTZ levels of theory. The results are shown in Figure V.4.

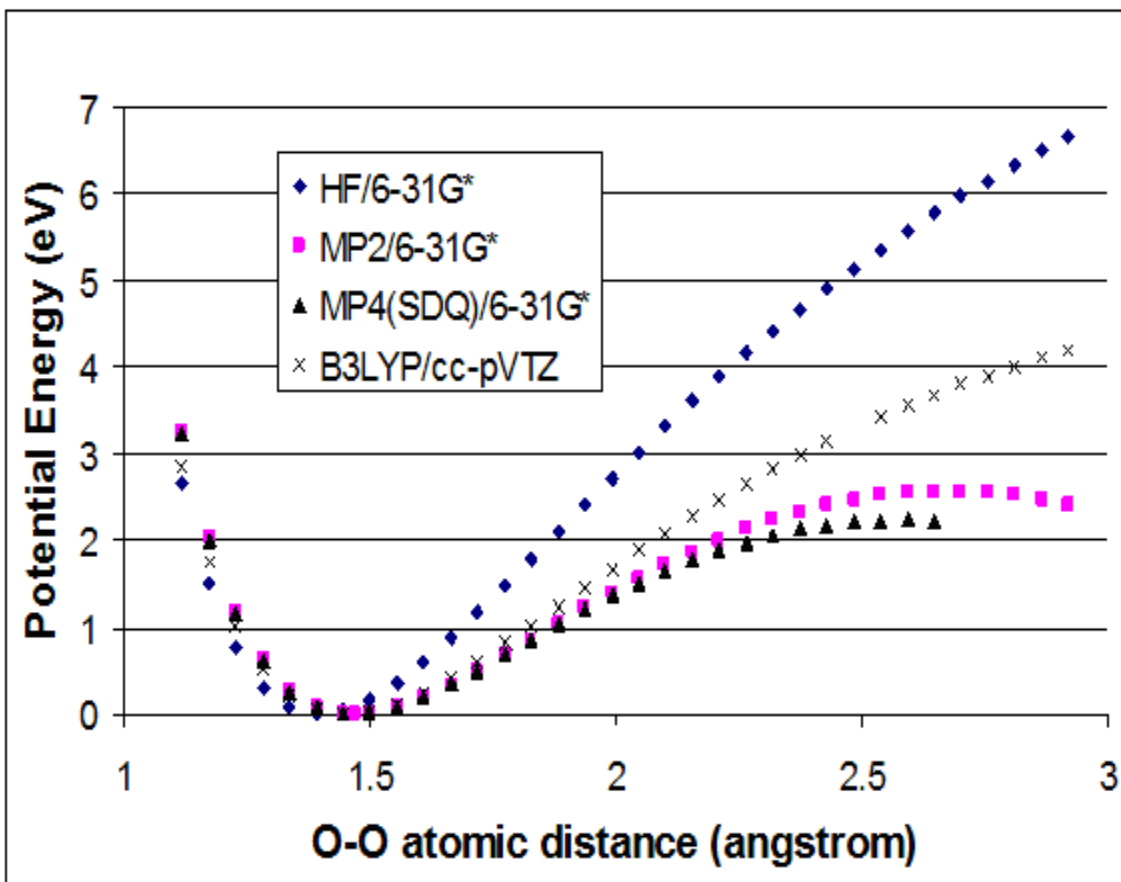


Figure V.4 The potential barrier investigated at HF/6-31G*, MP2/6-31G*, MP4/6-31G*, and DFT/B3LYP-cc-pVTZ levels of electronic structure theory.

In each calculation, the O-O bond distance is fixed at a certain value (from 1.12 to 2.92 Å), and the potential energy is optimized. According to MP2 and MP4(SDQ) investigation, O-O bond is believed to dissociate at 2.65 Å. At this distance, the potential height resulting from MP2 calculations is 2.57 eV. Therefore, the total energy (including zero point energy) used in our MD study ranges from 3.4 to 4.2 eV.

We also see a good agreement between MP2 and MP4(SDQ) calculations in term of O-O dissociating distance. However, the barrier height of MP4(SDQ) curve is 0.34 eV lower, which is about 2.24 eV. When simple restricted Hartree-Fock calculations are performed, no proof of dissociation is observed as the potential energy rises to 5.5 eV when O-O bond distance reaches 2.65 Å. DFT (B3LYP) calculations are executed on a large basis set (cc-pVTZ). The result clearly shows that there is no dissociation if O-O bond is less than 2.9 Å. In fact, based on a previous investigation, Kuhn *et al.*¹⁵ suggested that the bond will not be broken until 3.2 Å by executing second order perturbation complete active space calculations (CASPT2) and DFT calculations.

To investigate the O-O dissociation, quasi-classical dynamics is employed. Initially, HOOH is assigned at its equilibrium configuration, and vibrational energy is introduced into each mode using the projection method.¹⁰⁸ The trajectory is then integrated for a randomized period of time with no angular momentum. Subsequently, excitation energy is introduced into the six vibrational modes equally. At this point, it is guaranteed that HOOH has a configuration with randomized geometric coordinates and momenta at a certain energy level. The trajectory is then integrated with a fixed step size of 1.018×10^{-16} s using the fourth-order Runge-Kutta integration method.¹⁰⁹ A trajectory is terminated if 5 ps elapse or O-O bond rupture is found (the distance reaches 2.65 Å), and trajectory time is recorded. The rate coefficient k at a given total energy is calculated based on the first-order decay plot of 1,000 sample trajectories. In Figure V.5, a first-order typical decay plot at 3.4 eV of total energy is shown. Table V.5 gives calculated rate coefficients at the investigated total energy levels.

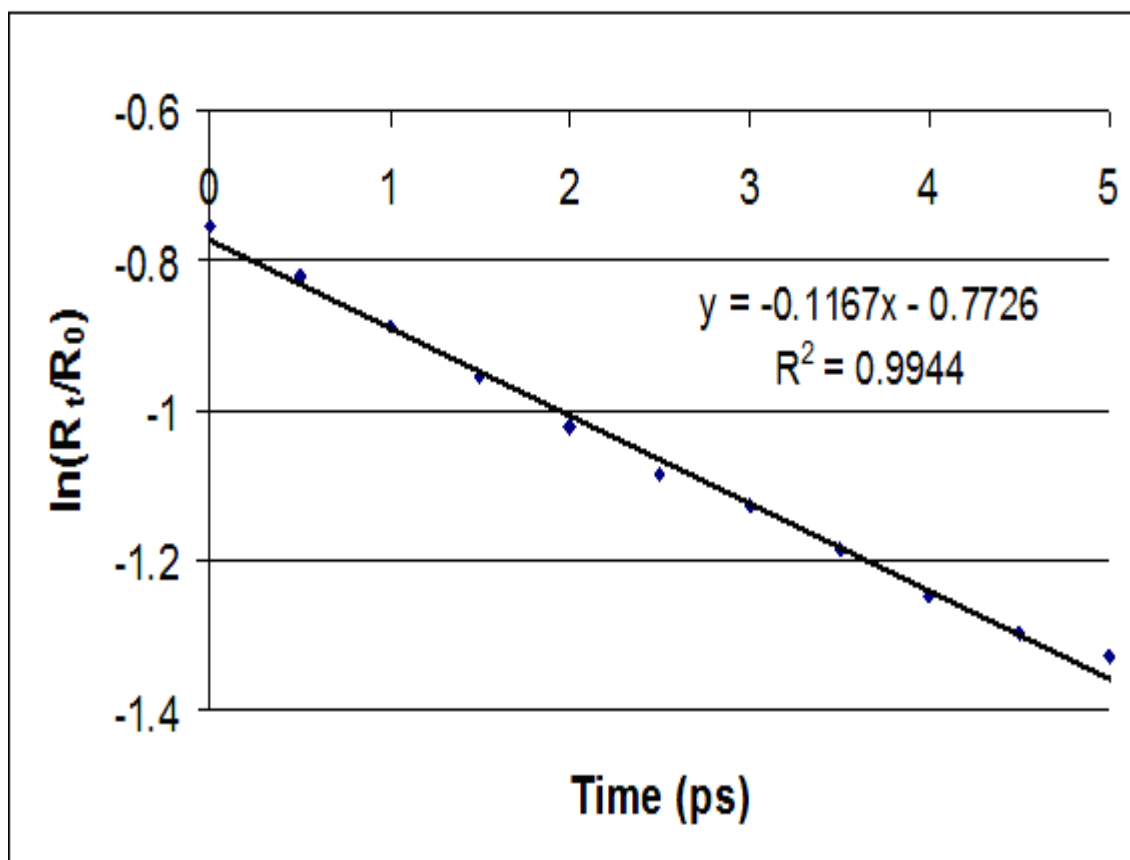


Figure V.5 First-order decay plot of O-O dissociation reaction when all vibrational modes are equally excited and the total energy equals 3.4 eV (including zero point energy). The excellent linearity of the results shows that we have very good statistical accuracy and no angular momentum present in HOOH.

Table V.5 First-order rate coefficients obtained from decay plots at five different total energies. The reported standard deviations are obtained from statistical analysis of the linear least-squares fit to the data.

Total energy (eV)	Rate (ps ⁻¹)	Standard deviation (ps ⁻¹)
3.4	0.117	0.003
3.6	0.160	0.004
3.8	0.205	0.005
4.0	0.281	0.007
4.2	0.324	0.009

In this section, we also present a proof of convergence, which means we have a sufficient amount of data points to describe the PES. A very important criterion to testify the convergence of a database is the consistency of MD calculations. If the rate coefficient at a certain energy level does not change significantly when more configurations are added to the database, a convergence can be concurred. When 25,608 points (100% of our sampled data points) are used to train a NN, the rate coefficient obtained at 3.4 eV of internal energy is 0.117 ± 0.002 ps⁻¹. Several NNs are also trained using less data points, and the rate coefficients at 3.4 eV are calculated on each surface, and the results of which are given in Table V.6.

Table V.6 First-order dissociation rate coefficients for HOOH at an internal energy of 3.4 eV using NN surfaces trained with smaller data randomly selected from the total database.

Number of points in the database	% of the total	Rate (ps ⁻¹)
20,486	80	0.094 ± 0.002
23,048	90	0.126 ± 0.002
24,328	95	0.121 ± 0.002
25,608	100	0.117 ± 0.002

When 90% of the database is used and the result is compared to the 100% case, the difference between the two rate coefficients is 0.009 ps⁻¹. In the 95% case, the rate is further converged by a difference of 0.004 ps⁻¹. This proves an essential convergence of the database, and we expect no significant changes occur when more data points are sampled and added.

The rate coefficients given in Table V.5 are fitted to the classical Rice-Ramsperger-Kassel (RRK) equation. According to RRK theory, rate k is related to the investigated internal energy by the following equation

$$k = f \left(1 - \frac{E_0}{E}\right)^{s-1} \quad (24)$$

where E_0 is the dissociation barrier height. According to our MP2/6-31G* calculations, E_0 is reported as 2.57 eV. In Figure V.6, the variation of $\ln(k)$ is linear with $\ln(1-E_0/E)$. Although the system is studied under quasi-classical aspect, the linearity of RRK plot

clearly shows MD trajectories behave similar to a pure classical case. This is not a surprising consequence because the zero-point energy is allowed to spread over all vibrational modes for a random period of time before the insertion of excitation energy using the projection method.

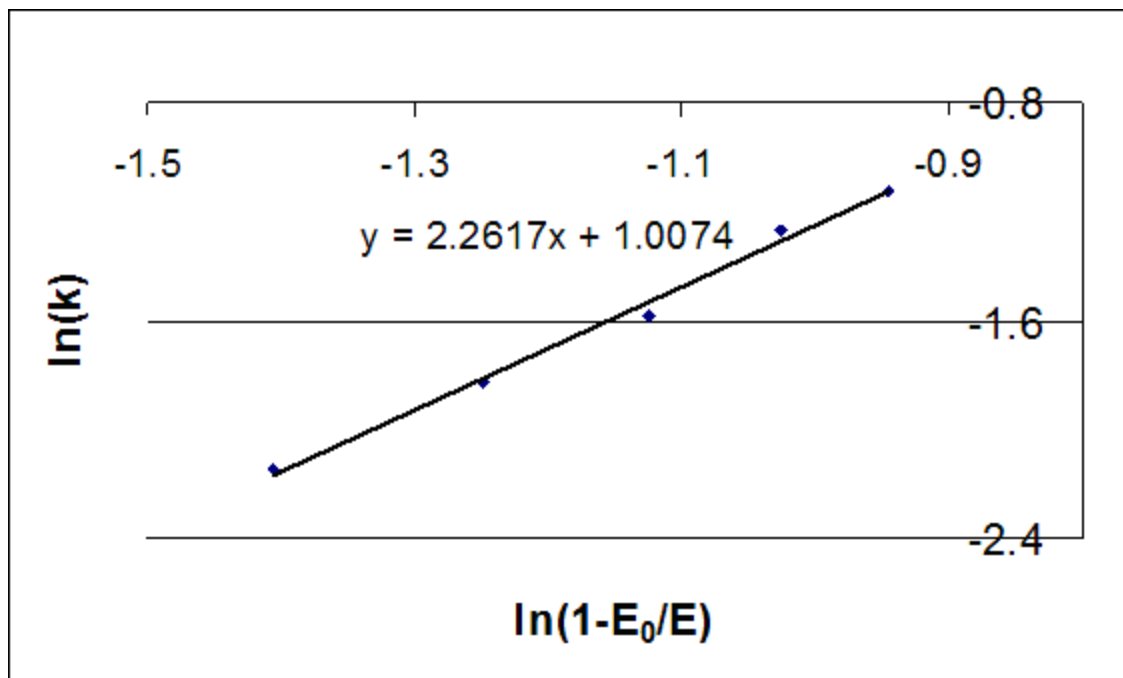


Figure V.6 RRK plot of O-O dissociation rate coefficients.

From the RRK plot, we obtain an s value of 3.26. Theoretically, s represents the number of effective modes during a MD study. With a value of 3.26, we conclude that there are three ineffective modes during the dissociation of O-O bond, which include OH symmetric and antisymmetric stretching modes, and the torsional mode. We also obtain an f value of 9.58 ps^{-1} , which corresponds to a wave number of 320 cm^{-1} . This energy is relatively low, and in fact, provides evidence of slow dissociation of O-O bond due to internal hydrogen bonding.

The O-O dissociation dynamics observed in this study suggests that internal hydrogen bonds tend to form during the dissociating process. When O-O distance is stretched, one or both of the hydrogen atoms tend to move closer to the other oxygen atom. As this occurs, one or both of the bending angles decrease. As a result, the dissociation of O-O bond happens slower than expected due to the barrier height and O-O stretching frequency. An illustrative snapshot is presented in Figure V.8 that shows how hydrogen bonds are formed during the dissociating process, which was also suggested by Harding in a previous study.⁶⁴

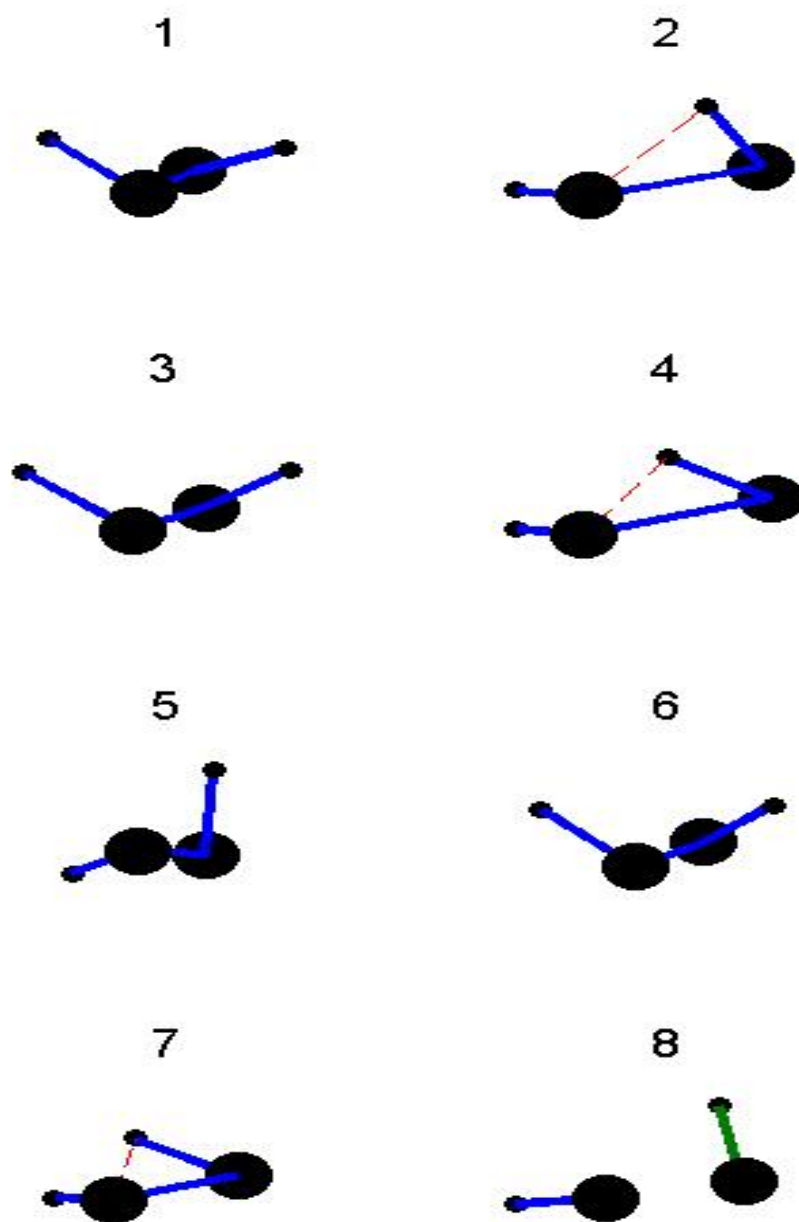


Figure V.8 Snapshots of HOOH configurations occurring during a trajectory resulting in O-O bond rupture that illustrate the effect of hydrogen bonding on the dissociation

processes. In snapshot (1), HOOH configuration is near equilibrium as the excitation energy is just added. In snapshots (2)-(4), the O-O bond distance increases from 1.46 Å to 1.85 Å. In snapshot (4), the H-O-O angle has decreased to about 60° to form a hydrogen bond. This interaction prevents O-O bond rupture, and the O-O distance begins to drop to 1.37 Å shown in snapshot (5). The sequence is repeated again in snapshots (6) and (7), but this time, the hydrogen-bond interaction is not sufficiently strong to prevent the O-O bond rupture shown in snapshot (8).

V. Summary

The O-O bond dissociation of hydrogen peroxide has been investigated on a PES obtained by fitting 25,608 points of MP2/6-31G* energies using a committee of five NNs, each of which has 34 neurons in the hidden layer. The average absolute error and root mean squared error are 0.0060 eV (0.58 kJ mol⁻¹) and 0.0099 eV (0.96 kJ mol⁻¹), respectively (the maximum potential energy is 4.36 eV).

We have executed a fit using the SVM method, and this new fit is compared to the previous NN fit. With the same set of data, the SVM fit using a radial basis kernel yields 447 support vectors, giving a total number of parameters to be 11.5 times higher than the total number of NN parameters (a single NN). Therefore, it consumes more computational time to execute a MD trajectory on the SVM surface (about 12 times). The average absolute error for the SVM function is 0.0311 eV. These results indicate the SVM method is not suitable MD studies, although it has been proven to be powerful in data classification. Other suitable methods for fitting data that have been introduced are NN,^{29-32,74-79,81,85,86} IMLS,²¹ and the Shepard interpolation method.²⁰

The initial set of HOOH configurations is obtained by performing partial optimization of potential energy. In each optimization, the dihedral angle and one of the five valence internal coordinates are assigned at some constant values. After constructing the initial set, more data points are obtained using the new sampling method. This method allows us to sample HOOH configurations without using MD trajectories. In this method, we analyze the data and find the regions that lack of data points. To do this, a derivative test is executed by comparing the NN derivatives with respect to the inputs to the real MP2 derivatives. If the difference is more than 3%, the region is identified, and more data points are generated in that region with two restricted conditions: the potential energy does not exceed 4.36 eV and the distance of a new point is greater than a defined threshold. Once identified, *ab initio* calculations are executed for the new point, and if the difference between the *ab initio* energy and the NN-predicted energy (temporary fit) is greater than 1%, the point is added to the database. This procedure allows us to sample HOOH configurations in a six-dimensional hyperspace with a more uniform density, thereby enhancing the NN fitting accuracy.

After four iterations, a database of 12,804 points is produced, and latter duplicated due to the symmetry of HOOH, which gives us a set of 25,608 configurations. Normal mode analysis on the NN surface is performed, and gives very good agreement with MP2-predicted wave numbers.

The present sampling method is based on an analysis of gradients; therefore, our NN committee is in essence of fitting both potential energies and gradients. This is reflected by a good agreement between NN-predicted and *ab initio* wave numbers. It is also supported by the low fitting error of the testing set (0.0060 eV).

The dissociation of O-O bond is investigated at 3.4, 3.6, 3.8, 4.0, and 4.2 eV of internal energy. The excitation energy is distributed equally among the six vibrational modes, then the trajectories are integrated using the Runge-Kutta method.¹⁰⁶ On an average, it requires about 45 s of computer time to execute a full trajectory (lasting for 5 ps in molecular time) on the NN committee surface when a 2.4 GHz clock-speed CPU is employed.

From MD trajectories, we are able to calculate the first-order rate coefficients at different energy levels with small deviations. A RRK plot is made that describes the relationship between internal energies and reaction rate constants. The plot exhibits good linearity, and can be used to predict the rate coefficients at various energy levels that are not reported in this study.

From RRK theory, we suggest that three of the six vibrational modes are not effective during the dissociation process. Although we execute quasi-classical trajectories in this study, the linearity of RRK plot shows that the dynamics of HOOH behaves classically because the zero-point energy is allowed to spread over the six vibrational modes for a random period of time.

Internal hydrogen bonding occurs during the process. We monitor the dissociation, and observe the tendency of a hydrogen atom moving toward the oxygen that it does not chemically bond to. This behavior gives the molecule more stability, and makes the reaction rate slower than expected.

CHAPTER VI

CONCLUSION

We have successfully employed the NN method to construct PES's for three chemical systems with different levels of complexity. In each study, we present a modified procedure and contribute a new approach to study a molecular system. The new approach may be a new data sampling method, or an improvement of the fitting technique. Besides NN fitting, SVM fitting is also presented and tested in the last study of HOOH, and comparisons are made to the NN method.

In our first study, an analytic PES of HONO is developed using the traditional sampling method, which is known as "novelty sampling."³¹ More than 21,000 configurations are obtained during the sampling process, and the data represents *cis* and *trans* configurations of HONO as well as the dissociation pathway of N-O bond. Subsequently, MD trajectories are performed on the surface. The resulting rate coefficients in *cis-trans* isomerizations when applying different types of excitations do not differ a lot. When we compare our results to the results in a previous study conducting MD on an empirical surface, we observe that the intra-mode coupling on the NN *ab initio* surface is significantly larger. This is also the case for N-O dissociation. It is believed that the internal energy is transferred very rapidly among the vibrational modes on the NN surface, which results in the excitations to be somehow similar.

The PES fit of HONO only requires 41 neurons, which is a computational advantage compared to the other methods such as IMLS and Shepard interpolation. Therefore, the use of a committee of five NN's has been proposed, and the fitting error of the NN committee is less than the error of each individual NN because nearly random errors are mostly cancelled. Thus, in the latter two studies (BeH_3 ³² and HOOH ⁶³), we employ NN committees to fit the data.

A direct comparison between the NN and Shepard interpolation methods is presented in the second study. We conduct a study on the BeH_3 system which has been previously studied by Collins and Zhang.³³ In that study, 1,300 data points were sampled using classical dynamics, and *ab initio* energy and gradients with respect to atomic distances are computed using MP2/6-311G(d,p) level of theory. The fit was performed on those 1,300 points, and the reported testing error was 0.00424 eV or 0.41 kJ mol⁻¹. Employing the data contributed by Collins and Zhang,³³ we sample more configurations using the “novelty sampling” method³¹ and end up with 19,208 configurations in the final database. In this study, it is realized that the nature of the investigated chemical reaction also affects the number of size of fitting NN's. With large extension of atomic distances in hyperspace, we have to use 60 neurons in each individual NN to provide a good fit for BeH_3 . The fitting error of the committee (of five NNs) is 0.44 kJ mol⁻¹, which is very comparable to the Shepard interpolation error.

Since the fit in Shepard interpolation method is a summation of weighted Taylor expansions around a group of selected data points, the weights have to be re-calculated at every new point during trajectories, and the computational cost becomes huge. The NN committee, on the other hand, uses a number of fixed parameters and therefore, gives a

noticeable advantage in computational cost with a comparable fitting error to the other method (0.44 kJ mol^{-1} versus 0.41 kJ mol^{-1}).

The reaction probability is calculated with various translational energy levels from 0.415 eV to 0.829 eV and no total angular momentum. This method was also employed by Collins and Zhang in their investigation. Our maximum probability is 0.152 ± 0.005 , which is about 38% higher than the maximum probability reported by Collins and Zhang³³ (0.11). The reaction threshold is found by conducting classical dynamics with no zero-point energy at various translational levels. At 0.56 eV , products begin to occur, and this energy is concluded to be the reaction threshold within the statistical accuracy of the trajectory study. Our reaction threshold is in good agreement with the potential barrier reported by Collins and Bettens.³⁷

During the first two studies (HONO^{24} and BeH_3^{32}), approximately 20 thousands points are collected using the “novelty sampling” procedure proposed by Raff and co-workers.³¹ The complexities of these two systems are very distinct. HONO is considered to be more complicated in term of chemical intuition as it has two different reaction channels and two stationary configurations (*cis* and *trans*) as observed in our first study. Moreover, with 24 electrons, the electronic configuration of HONO is more complex than that of BeH_3 , which only has 7 electrons. On the other hand, the hyperspace of BeH_3 is much larger as we attempt to describe the collision between two separate molecules. It can be concluded that the size of hyperspace as well as the number of reaction channels determines the number of required hidden neurons to fit the PES.

The uniformity of data points is very important to a NN fit as we have observed during the studies of HONO^{24} and BeH_3 .³² If the density of data points in a particular

region is much higher than the others, the NN would automatically focus on the high density region, and tend to neglect the other regions. In the “novelty sampling” procedure,³¹ MD trajectories are employed to sample configurations. As a result, the low potential energy region tends to have low density of configurations, and thus, is not well-fitted by NN’s. In fact, this region is very important for MD studies because the initial molecular sampling process always requires to be done in this region. Therefore, we have introduced another method to improve the issue, which is referred to the “gradient sampling” method.

In the last study of HOOH,⁶³ we have successfully conducted a new sampling method, which not only sample a significant number of configurations to describe the chemical reaction, but also helps to improve the fitting accuracy. This method is based on searching for regions that are lacking data points in a defined region (<4.36 eV in the HOOH case particularly). Since the NN method fits most data very well, it is hard to identify the “bad” regions by just looking at the NN outputs. The output gradients, on the other hand, truly describe the shape of a function. If the fitted function is very distinct from the real function, the gradients resulted from the two functions will be much different (we may refer this to over-fitting). Adopting this idea, we are able to identify those “bad regions” by gradient analysis and sample more useful data points.

By sampling configurations without MD trajectories, we are able to collect data points more uniformly. Moreover, the requirement of well-fitted gradients also helps to improve the accuracy. The final fitting error when a NN-committee with five members is employed is reported as 0.0060 eV, one of the best fitting errors reported in the literature for a four-body system.

It is observed during the investigation of HOOH dynamics that internal hydrogen bonding is very likely to form, which resists the dissociation of O-O bond. Unless very high excitation energy is used, the dissociation will not happen. We also observe this interesting property, which was suggested in a previous study by Harding.⁶⁴

The success of NN method to describe PES's of a wide variety of chemical systems accurately really motivates us to extend our research to more complicated chemical reactions. Moreover, the use of a different system of coordination (beside the traditional Cartesian coordination) should be considered to use in molecular dynamics studies because NN's are capable of generalizing any types of mathematical functions. The change of coordination system is a promising approach because it may help either to reduce computational cost (by cutting down the number of differential equations in the classical Hamiltonian), or to enhance fitting accuracy. However, modifications of the kinetic terms in any molecular systems are not easy. In fact, it is difficult to modify the kinetic terms of a complex system that has many atoms. Therefore, the change of coordinate system has to be evaluated carefully, as we do not want to create too many mathematical problems just to reduce computational cost.

A test of the SVM method gives a good picture of its advantages and disadvantages in MD studies. Although the SVM method is very efficient and widely use in data classification, the ability to describe a mathematical function is still limited, especially when fitting gradients are also required to be accurate. The radial basis kernel in the SVM method is very similar to a feature in the NN method, which is known as radial basis network, and these two features are considered as not useful in chemical reaction dynamics when we consider their fitting accuracy and computational cost.

REFERENCES

- ¹ P. M. Morse, *Phys. Rev.* **34**, 57 (1929).
- ² M. P. Sudhakaran and L. M. Raff, *Chem. Phys.* **95**, 165 (1985).
- ³ Y. Guan and D. L. Thompson, *Chem. Phys.* **139**, 147 (1989).
- ⁴ Y. Guo and D. L. Thompson, *J. Chem. Phys.* **118**, 1673 (2003).
- ⁵ J. S. Wright and S. K. Shih, *J. Chem. Phys.* **73**, 5204 (1980).
- ⁶ H. B. Schlegel and J. J. McDouall, in *Computational Advances in Organic Chemistry*, Ed. C. Ogretir and I. G. Csizmadia (Kluwer Academic, The Netherlands, 1991) 167-185.
- ⁷ R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta.* **35**, 33 (1974).
- ⁸ R. J. Buenker, S. D. Peyerimhoff, and W. Butscher, *Mol. Phys.* **35**, 771 (1978).
- ⁹ D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.* **98**, 1358 (1993).
- ¹⁰ R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ¹¹ T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- ¹² K. A. Peterson, D. E. Woon, and T. H. Dunning Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- ¹³ A. Wilson, T. van Mourik, and T. H. Dunning Jr., *J. Mol. Struct. (Theochem)* **388**, 339 (1997).
- ¹⁴ A. Rahaman and L. M. Raff, *J. Phys. Chem. A* **105**, 2156 (2001).
- ¹⁵ B. Kuhn, T. R. Rizzo, D. Luckhaus, M. Quack, and M. A. Suhm, *J. Chem. Phys.* **111**, 2565 (1999).
- ¹⁶ A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).

- ¹⁷ C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- ¹⁸ K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990).
- ¹⁹ K. Andersson, P. A. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992).
- ²⁰ J. Ischtwan and M. A. Collins, *J. Chem. Phys.* **100**, 8080 (1994).
- ²¹ G. C. Maisuradze, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **119**, 10002 (2003).
- ²² M. T. Hagan, H. B. Demuth, and M. Beale, *Neural Network Design* (PWS, MA, 1996).
- ²³ C. Cortes and V. Vapnik, *Machine Learning* **20**, 273 (1995).
- ²⁴ H. M. Le and L. M. Raff, *J. Chem. Phys.* **128**, 194310 (2008).
- ²⁵ P. A. McDonald and J. S. Shirk, *J. Chem. Phys.* **77**, 2355 (1982).
- ²⁶ A. E. Shirk and J. S. Shirk, *Chem. Phys. Lett.* **97**, 549 (1983).
- ²⁷ L. Khriachtchev, J. Lundell, E. Isoniemi, and M. Rasanen, *J. Chem. Phys.* **113**, 4265 (2000).
- ²⁸ Y. Guan, G. C. Lynch, and D. L. Thompson, *J. Chem. Phys.* **87**, 6957 (1987).
- ²⁹ D. I. Doughan, L. M. Raff, M. G. Rockley, M. Hagan, P. M. Agrawal, and R. Komanduri, *J. Chem. Phys.* **124**, 054321 (2006).
- ³⁰ M. Malshe, L. M. Raff, M. G. Rockley, P. M. Agrawal, and R. Komanduri, *J. Chem. Phys.* **127**, 134105 (2007).
- ³¹ L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, *J. Chem. Phys.* **122**, 084104 (2005).
- ³² H. M. Le and L. M. Raff, *J. Phys. Chem.* (in submission).

- ³³ M. A. Collins and D. H. Zhang, *J. Chem. Phys.* **111**, 9924 (1999).
- ³⁴ D. H. Zhang and J. Z. H. Zhang, *J. Chem. Phys.* **101**, 1146 (1994).
- ³⁵ D. H. Zhang, J. C. Light, and S. Y. Lee, *J. Chem. Phys.* **109**, 79 (1998).
- ³⁶ D. H. Zhang and S. Y. Lee, *J. Chem. Phys.* **110**, 4435 (1998).
- ³⁷ M. A. Collins and R. P. A. Bettens, *Phys. Chem. Chem. Phys.* **1**, 939 (1999).
- ³⁸ J. Koput and K. A. Peterson, *J. Chem. Phys.* **125**, 044306 (2006).
- ³⁹ M. Head-Gordon, J. A. Pople, and M. J. Frisch, *Chem. Phys. Lett.* **153**, 503 (1988).
- ⁴⁰ M. J. Frisch, M. Head-Gordon, and J. A. Pople, *Chem. Phys. Lett.* **166**, 275 (1990).
- ⁴¹ M. J. Frisch, M. Head-Gordon, and J. A. Pople, *Chem. Phys. Lett.* **166**, 281 (1990).
- ⁴² M. Head-Gordon and T. Head-Gordon, *Chem. Phys. Lett.* **220**, 122 (1994).
- ⁴³ S. Saebo and J. Almlöf, *Chem. Phys. Lett.* **154**, 83 (1989).
- ⁴⁴ C. Moller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- ⁴⁵ J. Cizek, *Adv. Chem. Phys.* **14**, 35 (1969).
- ⁴⁶ G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- ⁴⁷ G. E. Scuseria, C. L. Janssen, and H. F. Schaefer III, *J. Chem. Phys.* **89**, 7382 (1988).
- ⁴⁸ G. E. Scuseria and H. F. Schaefer III, *J. Chem. Phys.* **90**, 3700 (1989).
- ⁴⁹ T. R. Rizzo, C. C. Hayden, and F. F. Crim, *Faraday Discuss. Chem. Soc.* **75**, 223 (1983).
- ⁵⁰ H. R. Dubal and F. F. Crim, *J. Chem. Phys.* **83**, 3863 (1985).
- ⁵¹ T. M. Ticich, T. R. Rizzo, H. R. Dubal, and F. F. Crim, *J. Chem. Phys.* **84**, 1508 (1986).
- ⁵² N. F. Scherer and A. H. Zewail, *J. Chem. Phys.* **87**, 97 (1987).

- ⁵³ T. M. Ticich, M. D. Likar, H. R. Dubal, L. J. Butler, and F. F. Crim, *J. Chem. Phys.* **87**, 5820 (1987).
- ⁵⁴ X. Luo and T. R. Rizzo, *J. Chem. Phys.* **93**, 8620 (1990).
- ⁵⁵ X. Luo and T. R. Rizzo, *J. Chem. Phys.* **94**, 889 (1991).
- ⁵⁶ X. Luo, P. Fleming, and T. R. Rizzo, *J. Chem. Phys.* **96**, 5659 (1992).
- ⁵⁷ M. Brouard and R. Mabbs, *Chem. Phys. Lett.* **204**, 543 (1993).
- ⁵⁸ B. Kuhn and T. R. Rizzo, *J. Chem. Phys.* **112**, 7461 (2000).
- ⁵⁹ S. Y. Lin and H. Guo, *J. Chem. Phys.* **119**, 5867 (2003).
- ⁶⁰ R. Chen, G. Ma, and H. Guo, *J. Chem. Phys.* **114**, 4763 (2001).
- ⁶¹ Y. Guo, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **121**, 5091 (2004).
- ⁶² G. G. Maisuradze, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **121**, 10329 (2004).
- ⁶³ H. M. Le, S. Huynh, and L. M. Raff, *J. Chem. Phys.* **131**, 014107 (2009).
- ⁶⁴ L. B. Harding, *J. Phys. Chem.* **95**, 8653 (1991).
- ⁶⁵ K. C. Thompson and M. A. Collins, *J. Chem. Soc., Faraday Trans.* **93**, 871 (1997).
- ⁶⁶ K. C. Thompson, M. J. T. Jordan, and M. A. Collins, *J. Chem. Phys.* **108**, 8302 (1998).
- ⁶⁷ R. P. A. Bettens and M. A. Collins, *J. Chem. Phys.* **111**, 816 (1999).
- ⁶⁸ M. A. Collins, *Theor. Chem. Acc.* **108**, 313 (2002).
- ⁶⁹ G. E. Moyano and M. A. Collins, *J. Chem. Phys.* **119**, 5510 (2003).
- ⁷⁰ Y. Guo, L. B. Harding, A. F. Wagner, M. Minkoff, and D. L. Thompson, *J. Chem. Phys.* **126**, 104105 (2007).

- ⁷¹ R. Dawes, D. L. Thompson, A. F. Wagner, and M. Minkoff, *J. Chem. Phys.* **128**, 084107 (2008).
- ⁷² M. Hornik, M. Stinchcombe, and H. White, *Neural Networks* **2**, 359 (1989).
- ⁷³ S. Lawrence and C. L. Giles, *International Joint Conference on Neural Networks* **1**, 1114 (2000).
- ⁷⁴ G. B. Orr, “Overfitting,” Willamette University, 31 Jul. 2009, <<http://www.willamette.edu/~gorr/classes/cs449/overfitting.html>>.
- ⁷⁵ T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, *J. Chem. Phys.* **103**, 5662 (1995).
- ⁷⁶ S. Manzhos and T. J. Carrington Jr., *J. Chem Phys.*, **125**, 084109 (2006).
- ⁷⁷ S. Manzhos and T. J. Carrington Jr., *J. Chem. Phys.* **125**, 194105 (2006).
- ⁷⁸ S. Manzhos and T. J. Carrington Jr., *J. Chem. Phys.* **127**, 014103 (2007).
- ⁷⁹ S. Manzhos and T. J. Carrington, Jr., *J. Chem. Phys.* **129**, 224104 (2008).
- ⁸⁰ P. M. Agrawal, M. Malshe, R. Narulkar, L. M. Raff, M. Hagan, S. Bukkapatnum, and R. Komanduri, *J. Phys. Chem. A* **113**, 869 (2009).
- ⁸¹ P. M. Agrawal, L. M. Raff, M. T. Hagan, and R. Komanduri, *J. Chem. Phys.* **124**, 134306 (2006).
- ⁸² J. Ludwig and D. Vlachos, *J. Chem. Phys.* **127**, 154716 (2007).
- ⁸³ P. M. Agrawal, A. N. A. Samadh, L. M. Raff, M. T. Hagan, S. T. Bukkapatnam, and R. Komanduri, *J. Chem. Phys.* **123**, 224711 (2005).
- ⁸⁴ T. B. Blank, S. D. Brown, *Anal. Chem. Acta.* **277**, 273, (1993).
- ⁸⁵ T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, *J. Chem. Phys.* **103**, 4129 (1995).

- ⁸⁶ F. V. Prudente, P. H. Acioli, J. J. S. Neto, *J. Chem. Phys.* **109**, 8801 (1998).
- ⁸⁷ A. Pukrittayakamee, M. Hagan, L. M. Raff, S. Bukkapatnam, and R. Komanduri, *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol. 17 (ANNIE2007), November, 2007.
- ⁸⁸ A. Pukrittayakamee, M. Malshe, M. Hagan, L. M. Raff, R. Narulkar, S. Bukkapatnam, and R. Komanduri, *J. Chem. Phys.*, **130**, 134101 (2009).
- ⁸⁹ O. Ivanciuc, *Rev. Comput. Chem.* **23**, 291 (2007).
- ⁹⁰ V. Vapnik and A. Lerner, *Automat. Remote Contr.* **24**, 774 (1963).
- ⁹¹ H. Drucker, C. J.C. Burges, L. Kaufman, A. Smola and V. Vapnik, *Advances in Neural Information Processing Systems* **9**, NIPS 1996, 155 (1997).
- ⁹² S. R. Gunn, *Support Vector Machine for Classification and Regression* (University of Southampton, England, 1998).
- ⁹³ M. J. Frisch, G. W. Trucks, H. B. Schlegel, GAUSSIAN 98, Revision A.1, Gaussian, Inc., Pittsburgh, PA, 2001.
- ⁹⁴ J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quant. Chem.* **14**, 545 (1978).
- ⁹⁵ R. J. Barlett and G. D. Purvis, *Int. J. Quant. Chem.* **14**, 516 (1978).
- ⁹⁶ R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971).
- ⁹⁷ W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- ⁹⁸ P. C. Hariharan and J. A. Pople, *Mol. Phys.* **27**, 209 (1974).
- ⁹⁹ M. S. Gordon, *Chem. Phys. Lett.* **76**, 163 (1980).
- ¹⁰⁰ G. A. Petterson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, *J. Chem. Phys.* **89**, 2193 (1988).

- ¹⁰¹ A. D. McLean and G. S. Chandler, *J. Chem. Phys.* **72**, 5639 (1980).
- ¹⁰² R. Krishnan, J. S. Binkley, S. Seeger, and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- ¹⁰³ H. W. Schranz, S. Nordholm, and G. Nyman, *J. Chem. Phys.* **94**, 1487 (1991).
- ¹⁰⁴ A. P. Cox, H. Brittain, and D. J. Finnigan, *Trans. Faraday Society* **67**, 2179 (1971).
- ¹⁰⁵ G. Herzberg, *Spectra of Diatomic Molecules* (Van Nostrand, NY, 1950).
- ¹⁰⁶ E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations: the Theory of Infrared and Raman Vibrational Spectra* (McGraw-Hill, NY, 1955).
- ¹⁰⁷ J. Murto, M. Rasanen, A. Aspiala, and T. Lotta, *J. Mol. Struct.* **122**, 213 (1985).
- ¹⁰⁸ L. M. Raff, *J. Chem. Phys.* **89**, 5680 (1988).
- ¹⁰⁹ L. M. Raff and D. L. Thompson, *Theory of Chemical Reaction Dynamics*, edited by M. Baer, Ed (CRC, Boca Raton, FL, 1985), Vol. III, p. 1.
- ¹¹⁰ F. Richter, M. Hochlaf, P. Rosmus, F. Gatti, and H. D. Meyer, *J. Chem. Phys.* **120**, 1306 (2003).

VITA

Hung M. Le

Candidate for the Degree of

Doctor of Philosophy

Thesis: MOLECULAR DYNAMICS STUDIES ON NEURAL NETWORK *AB INITIO*
POTENTIAL ENERGY

Major Field: Chemistry

Biographical:

Personal Data: DOB 01/09/1983, Nationality: Vietnamese.

Education: graduated from High school for the Gifted in Ho Chi Minh city,
Vietnam, received a B.S. in chemistry from the University of Central
Oklahoma, Edmond, OK.

Experience: teaching assistant in the Chemistry department at Oklahoma State
University from 1/2006 to 12/2009.

Name: Hung M. Le

Date of Degree: December 2009

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: MOLECULAR DYNAMICS STUDIES ON NEURAL NETWORK *AB INITIO* POTENTIAL ENERGY SURFACES

Pages in Study: 113

Candidate for the Degree of Doctor of Philosophy

Major Field: Chemistry

Scope and Method of Study: The use of NN methods in potential energy surface developments of three different systems: HONO, $BeH + H_2 \rightarrow BeH_2 + H$, and HOOH.

Findings and Conclusions: The neural network method has been employed to construct three analytic *ab initio* potential energy surfaces for three different chemical reactions, which are nitrous acid (HONO), $BeH + H_2$, and hydrogen peroxide (HOOH). Molecular dynamics studies are then executed on each surface to investigate the chemical reaction. Two different sampling techniques are used to sample data: novelty sampling and gradient sampling. These two techniques have been successfully used to sample configurations for the investigated molecular systems. Once a sufficient number of configurations is collected, the potential energy surface is constructed, and classical molecular dynamics can be easily utilized to simulate the chemical reactions in gas phase. From these studies, the neural network method is concluded to be a very promising method in theoretical reaction dynamics investigations because of its computational advantage and excellent fitting accuracy.

ADVISER'S APPROVAL: Lionel M. Raff
