A NEURAL NETWORK AND MOLECULAR DYNAMICS APPROACH FOR EVENT PROBABILITY PREDICTION DURING CHEMICAL VAPOR DEPOSITION OF A CARBON ATOM ON DIAMOND (111) SURFACE

By

SIVAKUMAR MATHANAGOPALAN

Bachelor of Science

Madurai Kamaraj University

Tamilnadu, India

2002

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
In partial fulfillment of
The requirements for
The Degree of
MASTER OF SCIENCE
July, 2005

A NEURAL NETWORK AND MOLECULAR DYNAMICS APPROACH FOR EVENT PROBABILITY PREDICTION DURING CHEMICAL VAPOR DEPOSITION OF A CARBON ATOM ON DIAMOND (111) SURFACE

Thesis Approved:

Dr. Ranga Komanduri

Thesis Adviser

Dr. Lionel. M . Raff

Dr. Martin. T. Hagan

Dr. Hongbin Lu

A. Gordon Emslie

Dean of the Graduate College

SUMMARY

Diamond synthesized by chemical vapor deposition (CVD) has vast applications in electronic devices, cutting tools, sensors, and micro-electro-mechanical devices (MEMS) because of its outstanding mechanical properties. A few of them being high hardness, high wear resistance, high thermal conductivity and low compressibility. Carbon atoms form one of the primary growth species in the synthesis of diamond. Molecular dynamics (MD) simulation is a tool used to study this growth mechanism at the atomic scale. However, MD simulations are time consuming and are bound to have statistical errors. In this study, molecular dynamics (MD) simulation was used to investigate the event probabilities of chemisorption, scattering, and desorption that occur when a carbon atom reacts with a hydrogenated diamond (111) surface at a radical site. The probabilities were calculated from 50 MD simulation trajectories. The input conditions namely, incident angle (θ), rotation angle (Φ), impact parameter (b), and kinetic energy of carbon atom (K) were selected from distribution functions. The inputs (θ, Φ, b, K) and the outputs (event probabilities) from MD simulations were then used to train multilayer neural networks (NN) to predict the aforesaid probabilities. To predict each of the three probabilities, 50 NN were trained. The probabilities obtained were the average of 50 neural networks. It was found that the NN predictions lie within one sigma standard deviation limit of MD. Also, the outputs of the neural networks were found to have less statistical fluctuation than MD. The NN were

also advantageous in studying the effect of any input parameter on the other, in minutes as compared to a number of days by MD. It was also found that, as impact parameter increases, chemisorption and desorption probabilities decrease whereas the scattering probability increases. As incident angle increases, chemisorption probability and desorption probabilities decrease, whereas scattering probability increases. As kinetic energy of the carbon atom increases, chemisorption probability increases, whereas scattering and desorption probabilities decrease. As rotational angle increases, there seems to be no effect on any of the three probabilities.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents and sister for their encouragement, support and patience all through this work.

I would like to thank the National Science Foundation (NSF) for the support of this work.

I would like to thank my advisor, Dr. Ranga Komanduri for providing me with the required resources and opportunity to pursue this research. I would also like to thank him for his support, guidance, and motivation.

I would like to extend my sincere gratitude to Dr. L.M. Raff, Dr. P.M. Agrawal and Dr. M. Hagan for educating me with the basic concepts of molecular dynamics and neural networks. Without their guidance and valuable suggestions, this work would not have been possible. I would like to thank Dr. H.B. Lu for agreeing to serve on my thesis committee.

I would also like to thank my friends, Bala and Abdul, for their many useful discussions and suggestions. I would also like to sincerely thank everyone from our molecular dynamics research group for their valuable comments towards accomplishing this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTD…)

TABLE OF CONTENTS (CONTD…)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (CONTD…)

CHAPTER 1

INTRODUCTION

Chemical vapor deposition (CVD) of diamond involves the reaction between the heated substrate and the gaseous precursors at the substrate temperature in a vacuum environment. This type of atomistic deposition can provide uniform and pure coatings [1]. As these coatings are used for a variety of applications, such as for wear resistance and corrosion resistance [2, 3], it is necessary that the desired film thickness and properties are obtained. However, there is no known experimental method to study this phenomenon at nanometric or atomistic level. Therefore, simulation tools become very useful in studying the phenomenon of chemical vapor deposition at the atomistic level.

Molecular dynamics simulations are the most widely used tool to study chemical reactions at the atomistic level. In this method, the motion of every atom in a system is followed by solving the Newtonian equations of motion over a very short period of time, on the order of 1 ps. In this study, molecular dynamics simulations have been used to study the deposition of carbon atoms on a diamond (111) substrate. Carbon atoms are one of the precursors in the formation of diamond - like carbon films for use in a wide variety of applications

ranging from wear resistant coatings to coatings of high thermal conductivity [4-6].

The main interest of the molecular dynamics researchers over the years has been to study the effect of parameters affecting the reaction of the precursor atoms on the substrate. The main constraint limiting this research is the immense amount of computer time involved [7-10] and the statistical errors present in the results.

Neural networks (NN), which work similar to neurons in a human brain, have found a variety of applications from banking to aerospace for problem solving. In this study, an effort has been made to use neural networks as a tool to predict the reaction probabilities for chemisorption, scattering, and desorption of carbon atoms incident on a diamond (111) surface. The neural networks developed in this study are significantly faster in predicting the event probabilities compared to MD simulations and have less statistical fluctuation.

This thesis is organized in the following order. After a brief introduction (Chapter 1), Chapter 2 presents a review of literature in the area of molecular dynamics simulations of CVD processes and NN applications. Chapter 3 presents the principles of molecular dynamics simulations. Chapter 4 describes the structure and type of neural networks that were used in this study. Chapter 5 gives the problem statement and the intended goals of this investigation. Chapter 6 describes the procedure used for the molecular dynamics simulations on the interaction of a carbon atom with the diamond (111) surface and also the neural networks implementation. Chapter 7 presents the results followed by a

discussion of the results. Chapter 8 provides conclusions drawn out of this work and future work.

CHAPTER 2

LITERATURE REVIEW

Diamond coatings have found a wide variety of applications from cutting tools to micro-electro–mechanical systems (MEMS) because of their outstanding mechanical properties, namely, high wear resistance and high temperature resistance [2,3]. One of the means by which diamond coatings have been synthesized is chemical vapor deposition (CVD). CVD involves the reaction between a heated substrate and reactive gases, called the precursors, to produce a desired film of a given thickness. There have been a wide variety of precursor gases used ranging from metals, metal hydrides, halides, halo hydrides, and metal organic compounds [1] in the CVD process. Thicknesses as low as several nanometers and as high as several micrometers can be obtained [1].

The diamond films formed by CVD are expected to be very precise in thickness [1] and should have the desired properties, high hardness, high wear and chemical resistance, to name a few. The type of film formed depends upon the parameters, such as the precursor used, aiming angle, rotation angle, and temperature of the substrate. The effect of most of these parameters, if not all, cannot be studied by experimental methods. Simulation is the best known available tool to investigate the effect of input parameters in a CVD process. As

the CVD coatings are deposited at the nanometric scale per second, molecular dynamics simulation is widely used to study the synthesis of diamond at this level.

2.1 Molecular dynamics simulation of chemical vapor deposition

Alfonso *et al.* [11, 12] studied the effects of hydrocarbon adsorption on diamond (100) flat and stepped surfaces by molecular dynamics simulations. They studied the adsorption of such hydrocarbons as CH_3, CH_2, C_2H, and C_2H_2 on a hydrogenated (100) diamond surface with a radical site. The empirical potential developed by Brenner *et al*. [14], together with a dynamical quenching technique, was employed in the studies. The effect of the binding energies, bond distances, bond angles, height of the adsorbed hydrocarbon from the surface, and tilt of bonds from the surface were investigated. It was found that, there was insignificant change in the binding energies of hydrocarbons adsorbed on flat and stepped surfaces.

Alfonso *et al.* [11, 12] also studied the adsorption probability for incident angles 0 ° and 45° of a CH_3 molecule and at substrate temperatures in the range of 300 – 900 K of 75 % and 50% hydrogenated diamond (100) surfaces. They conducted probabilistic studies based on 150 trajectories and monitored the reaction for each trajectory for a period of 2.5 ps. They found that the adsorption probability was found to be higher for normal incidence to the surface, and was found to increase with higher incident energies. They also concluded that a radical site is required for the CH_3 molecule to bind to the surface.

Du *et al.* [13] investigated the collision of C_{36} molecule on diamond (001) surface using MD simulations with the semi-empirical Brenner potential [14]. They studied the effect of varying the incident energy from 20 to 150 eV of C_{36} on its structure as well as the structure of the substrate. They also examined the threshold energy range within which C_{36} is chemisorbed on the substrate of diamond of different orientations. The real simulations of the trajectories were on the order of 2.5 to 3 ps. They found that the energy threshold for chemisorption is sensitive to the orientation of the incident C_{36} molecule and its impact position.

Gernster *et al.* [8] studied the effect of carbon deposition on diamond (111) substrate for diamond film formation averaging over the impact parameter for different incident angles (0°, 45°, 60°) of the C atom to the normal, different incident energies (25, 40, 50, 65 eV) of the C atom and for substrate temperatures of 100 K and 300 K. The time step used in these calculations was on the order 0.013 to 0.052 fs. They found that the formation of tetrahedral structure is favored by low incident energy species and low substrate temperature.

Neyts *et al.* [9] have studied the formation of diamond films using low energy hydrocarbon radicals (< 2 eV). The MD simulations were carried out on a previously simulated diamond film consisting of 712 atoms. The temperature of the substrate was maintained at 523 K. A time step of 0.5 fs was used. Each trajectory was carried out for 0.75 ps. The simulations investigated two conditions, namely,

- Low acetylene influx

- High acetylene influx

The sticking and desorption efficiencies of various growth species were studied. C_2 and C_2H molecules were found to be the most important species aiding diamond growth.

Huang *et al.* [10] studied the chemisorption of hydrocarbon molecules, namely, CH_3 and CH_2 on a hydrogen terminated diamond (001) surface using the many- body Brenner potential [14]. CH_3 and CH_2 hydrocarbons were made to impact on a hydrogenated and a non-hydrogenated substrate. The substrate was maintained at a temperature of 300 K. The hydrocarbons were projected on to the substrate at normal angle of incidence. They considered two cases. The first case considered was a mixture containing 60% CH_3 and 40% CH_2. The time between impingement of each hydrocarbon was set at 4 ps, in this case. The second case considered was a mixture of 98:2 hydrogen/methane gases which had the same amount of hydrocarbons as the previous one. A shower of hydrogen atoms (50 atoms) were made to impact on the diamond surface between hydrocarbon impacts to simulate a hydrogen terminated diamond surface and the shower lasted for 2 ps. Threshold energies for chemisorption of the hydrocarbons for fixed orientation and varying impact sites were determined.

Huang *et al.* [15] studied the energy threshold for chemisorption and variation of chemisorption with incident energy considering C_2H_2 and CH_3 radicals on a diamond (001) substrate covered with and without hydrogen atoms at a temperature of 300 K. The simulations for each trajectory lasted for 3 ps.

They found that molecules CH_3, CH_2, and hydrogen atoms form the important species in diamond growth.

Kaukonen *et al.* [16] studied the deposition effects of carbon atoms on a diamond (100) substrate made up of 640 atoms. They studied the effect of incident beam energy (up to 150 eV), substrate temperature, deposition angle of the carbon atom and thermal coupling constant on carbon deposition. They studied these effects by depositing 640 atoms on a substrate made up of 640 atoms and by depositing 500 atoms on a substrate made up of 1280 atoms. They found that , decreasing the substrate temperature and increasing the thermal conductivity at high deposition energies ere found to favor diamond growth.

Neyts *et al.* [17] studied the sticking and hydrogen abstraction reactions of various species, namely, C_2, CH_2, C_3, linear C_3H, cyclic C_3H, and cyclic C_3H_2 on a diamond-like carbon substrate (DLC). They randomly oriented these hydrocarbon species above the substrate containing 830 atoms. The hydrocarbons were aimed at random, in the x-y plane of the substrate. They have studied the effect of incident energy on the sticking probability. To calculate the sticking probability they considered 500 trajectories at energies of 0.11 eV and 1 eV. The Brenner *et al.* potential [14] and a time step of 0.5 fs were employed. The simulations were run for 2.5 ps for hydrocarbon energies of 0.1 eV and for 1.25 ps for energies of 1 eV. It was found that C_2 molecule in efficient in abstracting H atoms from the diamond like carbon (DLC) substrate.

Perry *et al.* [18,19] studied the rate coefficients, event probabilities and desorption probabilities of reactions of C_2H_2, C_2H, CH_3, CH_2, C_2H_4, C_2H_3, C_3H,

and C_n(1-3) at 1250 K on a diamond (111) terrace and a diamond (111) ledge surfaces. The terrace surface consisted of 147 atoms. All surface atoms were capped with hydrogen atoms except one. The ledge surface consisted of 147 atoms and two radical sites. The range of incident angles of the hydrocarbon varied from 0 to 50.19° and the aiming point was chosen randomly within a radius of 1 A. The number of trajectories used for calculating the event probability for each event varied from 250 to 1000. They have reported a computational speed of 30 min per trajectory on an ALPHA 3000/ Model 300 Workstation. The simulation time for individual trajectories varied in the range of 0.10 - 0.87 ps. Chemisorption reactions were found to occur at a higher rate on the ledge surface than on the terrace surface.

Zhu *et al.* [20] studied the interaction between C_2H_2 incident molecule and diamond (001) surface. The impinging C_2H_2 has an incident energy ranging from 1 to 20 eV. The Brenner potential which was developed from a Tersoff potential with bond-order correction, was used in the study. The chemisorption probabilities were studied for 200 collision events of C_2H_2. Each simulation trajectory lasted for 3 ps. Six types of chemisorption configurations were observed from the study.

Hu *et al.* [21,41] studied the adsorption probabilities of ethylene (C_2H_4) clusters, adamantine ($C_{10}H_6$) molecules, and fullerene (C_{20}) molecules at various incident angles and kinetic energies at a temperature of 300 K. Impact of 20 clusters on the hydrogenated diamond (111) substrate was studied. Impact at incident angles of 0°, 15°, 45°, and 60° were considered. Each trajectory was run

for 3 ps with a time step of about 0.2 fs. Adhesion of these radical atoms was studied with respect to their incident angle and incident energy. It was found that the adsorption probabilities were found to decrease with increasing incident angles.

2.2 Application of Neural networks (NN)

Artificial neural networks or neural networks (NN) are information processing systems that work similar to the neurons in the human brain. The history of neural networks dates back to 1940 when Mcculloh and Pitts [22] showed that artificial neurons could be used to compute any logical or arithmetic function. From then on, neural networks have been used in a variety of applications, such as aerospace, automotive, banking, defense, robotics, and speech recognition, to name a few [23-27].

Neural networks are used to predict the underlying function between a set of inputs and the corresponding outputs. Neural networks learn by example and react to situations to which they are trained. When they are given a known set of inputs within the scope of their training, they can accurately predict the output.

Natale *et al.* [28] used neural networks to study the relationship between inputs and outputs in an Atmospheric pressure chemical vapor deposition (APCVD) process of silicon dioxide. The neural network was trained using data from experiments. Inputs to the neural network were gas flow rates, temperature of injectors, nitrogen curtain flows, chamber pressure, butterfly purge valve pressure, and thermocouple temperature. The outputs were the weight percentage of dopants, namely, boron, and phosphorous. The thickness of the

film formed was also predicted. Thus, the neural network aided in the prediction of the best combination of input conditions to obtain a desired film thickness of a desired composition. The trained neural network was able to predict the output within 1% to 10% of the experimental values.

Han *et al.* [29] used neural networks to predict the optimized operating conditions in a Plasma enhanced chemical vapor deposition (PECVD) process. The neural network was trained for inputs namely, substrate temperature, pressure, RF power, and nitrous oxide flow. The outputs for training were the deposition rate, permittivity, film stress, uniformity, silanol concentration, and water concentration. The trained network was then used to obtain optimized conditions to produce novel film properties, such as 100% uniformity, low permittivity, stress free, and impurity free concentration.

Geisler *et al.* [30] used a five layer feed forward neural network to predict the properties of a Si_3N_4 film synthesized by plasma enhanced CVD (PECVD). The neural network was trained for six input variables, namely, substrate temperature, chamber pressure, RF power, NH_3 flow, SiH_4 flow, and N_2 flow. The film's refractive index, the effective lifetime, and positive charge density formed the outputs. To train the network, 47 data points were used. Sensitivity studies were carried out on the output.

The use of neural networks in molecular dynamics simulations is in its early stages. For the most part, they have been used to develop potentials using electronic structure calculations. Hobday *et al.* [31] used neural networks to develop a potential energy surface for a C – H system. Inputs to the neural

network consisted of pairwise terms, such as, direction cosines, bond lengths, cut – off functions, first neighbor information, such as its bond lengths and torsional angles and second neighbor information such as the number and type of atoms. The output was the potential energy of an atom. The numerical experiments carried out indicates the possibility for the development of a potential energy surface (PES). The PES although slower than Brenner potential by 60% – 80%, it is still inexpensive compared to the *ab initio* calculations

Lorenz *et al.* [32] developed a potential energy surface using neural networks trained for *ab-initio* data. The sticking probability of H_2 on a Pd (100) surface is determined using molecular dynamics calculations on the potential energy surface developed by neural networks. Out of the 659 *ab-initio* energies obtained, 619 were used for training and 40 were used for testing. The potential energy surface developed using neural network was very accurate in predicting the sticking probability of H_2 on a Pd (100) surface as compared to analytical potential energy surface.

Raff *et al.* [33] used neural networks to interpolate potential energy surfaces for Si_5 and vinyl bromide. The neural networks were trained for data from *ab-initio* calculations. The potential energy surface they developed can also be used for Monte Carlo studies, gas-phase chemical reactions, nanometric cutting and nanotribology. The method has a better computational accuracy and advantages over the existing methods and it is also easy to implement.

Sumpter *et al.* [34] used neural network as a tool to study the energy during stretch, bending, and torsion of various bonds in H_2O_2 molecule.

Trajectory calculations were performed on a potential energy surface derived from *ab-initio* calculations. The neural network in this case was trained to predict the mode energies within a few seconds and with a percentage error of 1% to 12% as compared to molecular dynamics simulations.

This literature survey indicates that if data from MD simulations of reaction studies in a CVD process can be used to train a neural network, the network can predict the results for various input conditions in very short time compared to pure molecular dynamics simulation studies.

CHAPTER – 3

PRINCIPLES OF MOLECULAR DYNAMICS (MD) SIMULATION

Molecular dynamics (MD) simulations have been a very effective tool in nanometric simulation studies [38, 39, 43]. Using molecular dynamics simulations, the positions, velocities, and acceleration of every atom in the system can be obtained at different times by solving Newtonian equations of motion.

In a system that consists of many atoms, the initial positions are assigned to the atoms of the system depending upon the initial configuration. The new positions of the atoms at the next time step are obtained by solving the Newtonian equations of motion [38, 39]. If V is the potential function, then the force experienced by an atom i is given by

$$F = -\frac{dV}{dr_i}, \tag{3.1}$$

where, F is the force and r_i is the position vector of the i^{th} atom at time t. Therefore, the acceleration of the atom can be obtained by

$$m\frac{d^2r}{dt^2} = -\frac{dV}{dr_i}. \tag{3.2}$$

The velocity and new position can be obtained by solving

$$v = a\delta t + v_0, \tag{3.3}$$

where, v is the velocity at time $t+\delta t$, v_0 is the velocity of the atom at time t_0, a is the acceleration, and δt is the time step. The time step used in this study is 0.5 fs. Having the velocity, the new position of the atoms is given by

$$r = v\delta t + r_0, \tag{3.4}$$

where, r is the new position of the atoms and r_0 is the old position of the atoms.

3.1 Integration Method

There are many integration methods, such as Runge–Kutta, velocity verlet, leapfrog, Adams – Bashforth- Moulton fourth order to integrate the equations of motion. Consider an example of the the Gear predictor - corrector method [38, 39]. In this method, the position, velocity, and acceleration of every atom in the system are calculated for each time step. The time step is 0.5 fs. The position, velocity, and acceleration of atoms are calculated in two steps, the predictor step and the corrector step.

Let v be the velocity, a the acceleration, and r the position of a given atom at time step t. The position, velocity, acceleration, and rate of change of acceleration at the next time step t+δt is predicted to be r^p, v^p, a^p, b^p. The values are calculated at the predictor step.

$$r^p(t+\delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + \frac{1}{6}\delta t^3 b(t) + ..., \tag{3.5}$$

$$v^p(t+\delta t) = v(t) + \delta t a(t) + \frac{1}{2}\delta t^2 b(t) + ..., \tag{3.6}$$

$$a^p(t+\delta t) = a(t) + \delta t b(t) + ..., \tag{3.7}$$

$$b^p(t+\delta t) = b(t) + ... \tag{3.8}$$

The force experienced by an atom in the system is given by

$$F = -\frac{dV(r^p)}{dr^p}. \tag{3.9}$$

This force is obtained by substituting the position r^p of each atom from the predictor step on to the potential function used in the study.

Corrected acceleration is given by

$$a^c = \frac{F}{m}. \tag{3.10}$$

The acceleration a^c thus obtained is used to find the errors in the values of the accelerations found in the predictor step.

Error in the acceleration value is given by

$$\Delta a\ (t+\delta t) = a^c(t+\delta t) - a^p(t+\delta t). \tag{3.11}$$

The difference in acceleration is used to correct the predicted values of position, velocity, acceleration, and rate of change of acceleration. The corrected values are given by

$$r^c(t+\delta t) = r^p(t+\delta t) + c_0\, \Delta a\ (t+\delta t), \tag{3.12}$$

$$v^c(t+\delta t) = v^p(t+\delta t) + c_1\, \Delta a\ (t+\delta t), \tag{3.13}$$

$$a^c(t+\delta t) = a^p(t+\delta t) + c_2\, \Delta a\ (t+\delta t), \tag{3.14}$$

$$b^c(t+\delta t) = b^p(t+\delta t) + c_3\, \Delta a\ (t+\delta t), \tag{3.15}$$

where, $c_0 = 1/6$, $c_1 = 5/6$, $c_2 = 1$, $c_3 = 1/3$.

Eqns (3.12) – (3.15) give the corrected values of position, velocity, acceleration, and rate of change of acceleration.

3.2 Potential function

Empirical potential, in general, is an analytical function used to describe the potential energy between the atoms in a chemical system. It has parameters that can be modified to describe the properties of the given material. For example, the many body empirical potential V developed by Brenner *et al*. [14], takes the form

$$V = \sum_{i} \sum_{(j>i)} [V^R(r_{ij}) - \bar{b}_{ij} V^A(r_{ij})]. \tag{3.16}$$

Here, the functions V^R and V^A are pair-additive interactions that represent all interatomic repulsions and attractions from valence electrons. r_{ij} represents the distance between neighbor atoms *i* and *j*.

The empirical bond order function $\bar{b}_{ij}$ takes the form

$$\bar{b}_{ij} = \frac{1}{2}[b_{ij}^{\sigma-\pi} + b_{ji}^{\sigma-\pi}] + b_{ij}^{\pi}. \tag{3.17}$$

The values for the functions $b_{ij}^{\sigma-\pi}$ and $b_{ji}^{\sigma-\pi}$ depend on the local coordination and bond angles for atoms i and j, respectively. Function b_{ij}^{π} can be further written as

$$b_{ij}^{\pi} = \pi_{ij}^{RC} + b_{ij}^{DH}. \tag{3.18}$$

The repulsive function $V^R(r)$ and the attractive function $V^A(r)$ have the forms given by Eqns (3.19) and (3.20), respectively.

$$V^R(r) = f^c(r)(1 + Q/r) A e^{-\alpha r}. \tag{3.19}$$

$$V^A(r) = f^c(r) \sum_{n=1,3} B_n e^{-\beta_n r}. \tag{3.20}$$

The parameters for carbon- carbon pair terms used are given in Table 3.1

Table 3.1 Parameters for carbon – carbon pair terms

B1 = 12388.79197798 eV	β1 = 4.7204523127 $Å^{-1}$	Q = 0.3134602960833 Å
B2 = 17.56740646509 eV	β2 = 1.4332132499 $Å^{-1}$	A = 10953.544162170 eV
B3 = 30.71493208065 eV	β3 = 1.3826912506 $Å^{-1}$	α = 4.7465390606595 $Å^{-1}$
D_{min} = 1.7	D_{max} = 2.0	

The function $b_{ij}^{\sigma-\pi}$ from Eqn (3.17) takes the form

$$b_{ij}^{\sigma-\pi} = [1 + \sum_{k(\neq i,j)} f_{ik}^{c}(r_{ik}) G(\cos(\theta_{ijk})) e^{\lambda_{ijk}} + P_{ij}(N_i^C, N_i^H)]^{-1/2}. \quad (3.21)$$

The subscripts represent the atom identity. The function $f^{c}(r)$ refers to the interactions of the nearest neighbors only. Function P is a bicubic spline and quantities N_i^C and N_i^H represent the number of carbon and hydrogen atoms that are neighbors to atom i. N_i^C and N_i^H are defined by

$$N_i^C = \sum_{k(\neq i,j)}^{\text{carbon atoms}} f_{ik}^{c}(r_{ik}). \quad (3.22)$$

$$N_i^H = \sum_{l(\neq i,j)}^{\text{hydrogen atoms}} f_{il}^{c}(r_{il}). \quad (3.23)$$

For solid-state carbon, λ and function P are assigned a value of zero. The function $G(\cos(\theta_{ijk}))$ in Eqn (3.21) adjusts the contribution each neighbor makes to the empirical bond order.

N_i^t is the coordination of atom i. It takes the form

$$N_i^t = N_i^C + N_i^H. \tag{3.24}$$

The term π_{ij}^{RC} in Eqn (3.18) signifies the influence of radical energies and π bonds configuration of bond energies and is given by

$$\pi_{ij}^{RC} = F_{ij}(N_i^t, N_j^t, N_{ij}^{conj}), \tag{3.25}$$

where ,

$$N_{ij}^{conj} = 1 + [\sum_{k(\neq i,j)}^{\text{carbon}} f_{ik}^c(r_{ik}) F(X_{ik})]^2 + [\sum_{l(\neq i,j)}^{\text{carbon}} f_{jl}^c(r_{jl}) F(X_{jl})]^2, \tag{3.26}$$

$$\begin{aligned} &F(x_{ik}) = 1, && x_{ik} < 2, \\ &F(x_{ik}) = [1 + \cos(2\pi(x_{ik} - 2))]/2, && 2 < x_{ik} < 3, \\ &F(x_{ik}) = 0, && 3 < x_{ik}, \end{aligned} \tag{3.27}$$

$$x_{ik} = N_k^t - f_{ik}^c(r_{ik}). \tag{3.28}$$

The term,

$$x_{ik} = N_k^t - f_{ik}^c(r_{ik}). \tag{3.29}$$

The term b_{ij}^{DH} in Eqn (3.18) is represented as shown in Eqn (3.30)

$$b_{ij}^{DH} = T_{ij}(N_i^t, N_j^t, N_{ij}^{conj})[\sum_{k(\neq i,j)} \sum_{l(\neq i,j)} (1 - \cos^2(\Theta_{ijkl})) f_{ik}^c(r_{ik}) f_{jl}^c(r_{jl})], \tag{3.30}$$

where,

$$\Theta_{ijkl} = e_{jik} e_{ijl}. \tag{3.31}$$

Function T_{ij} takes a tricubic spline form and functions e_{jik} and e_{ijl} are unit vectors in direction R. Values of the carbon – carbon cubic spline, T in Eqn (3.30) are given in Table 3.2

Table 3.2: Values of carbon-carbon cubic spline, T

i	**J**	**k**	**T(i, j, k)**
2	2	1	-0.070280085
2	2	9	-0.00809675

Value of $f^c(r)$ is a switching function given by,

$$\begin{aligned} f_{ij}^c(r) &= 1, && r < D_{ij}^{\min}, \\ f_{ij}^c(r) &= \left[1+\cos\pi((r-D_{ij}^{\min})/(D_{ij}^{\max}-D_{ij}^{\min}))\right]/2, && D_{ij}^{\min} < r < D_{ij}^{\max}, \\ f_{ij}^c(r) &= 0, && r > D_{ij}^{\max}. \end{aligned} \tag{3.32}$$

CHAPTER - 4

PRINCIPLES OF NEURAL NETWORKS

The human brain is made up of approximately 10^{11} neurons [22] which communicate among each other to accomplish any human activity. Artificial neural networks (NN) are simple representations of these biological neurons. These neural networks are not as powerful as the biological neurons in performing complex functions, but they can be well trained to perform simple functions. This chapter provides a brief summary of the architecture of a simple neural network as well as steps involved in the development of neural networks used in the present study.

4.1 Biological neurons

Artificial neural networks have derived their inspiration from biological neurons present in the human brain. There are 10^4 connections for each neuron [22]. Therefore, the human brain consists of $10^{11} \times 10^4$ connections which is quite complex.

A neuron consists of three principal components.

a. Dendrites

b. Cell Body

c. Axon

The dendrites are fiber - like structures which carry electrical signals into the cell body.

The cell body stores and sums up these electrical signals. The axon is a long fiber that carries electrical signals from the cell body to other dendrites. Figure 4.1 shows a simple neuron

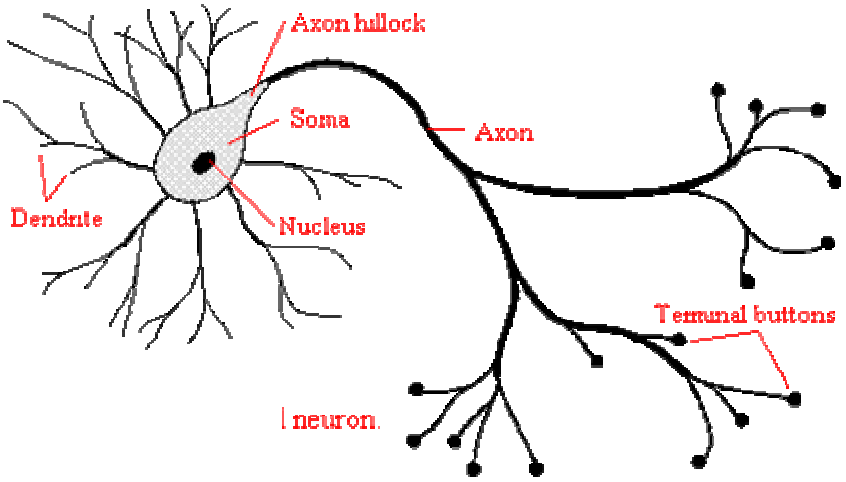


Figure 4.1 Schematic of a simple neuron [44]

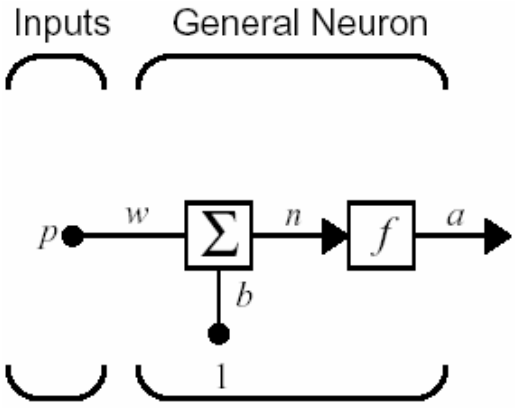
The point of contact of the axon of a neuron to the dendrite of another neuron is called the synapse. The arrangement of the neurons, the strength of the synapse and their communication determines the function of the biological neural network. The greater the strength of the synapse, the higher is the performance of the neural network.

4.2 Simple input neuron

Artificial neural networks are simple representations of biological neurons. A simple neural network is shown in Figure 4.2 [22], which can be mathematically represented as,

$$f(wp+b)=a. \tag{4.1}$$

A single input p is multiplied by a weight matrix w and the product wp is added to the product of the value 1 multiplied by a bias b. The total value $(wp+b)$ is called the net input and it is represented by n in Figure 4.2. This net input is then passed on to a transfer function f.



$$f(wp+b)=a$$

Figure 4.2 Single input neuron [22].

If the biological neuron and the single input neuron are to be compared, w would represent the strength of the synapse, the summation $(wp+b)$ would represent the cell body. The transfer function f and the output a would represent the signal carried away by the axon.

4.3 Multi layer network

Artificial neural networks have a wide variety of applications from banking to engineering [35-37]. These applications have more than one input and one or more outputs. A single input neuron, such as the one shown in Figure 4.2 cannot be used. Instead, multiple input neurons have to be implemented for such a situation. Multiple input neurons together with a multi- layer network are shown in Figure 4.3

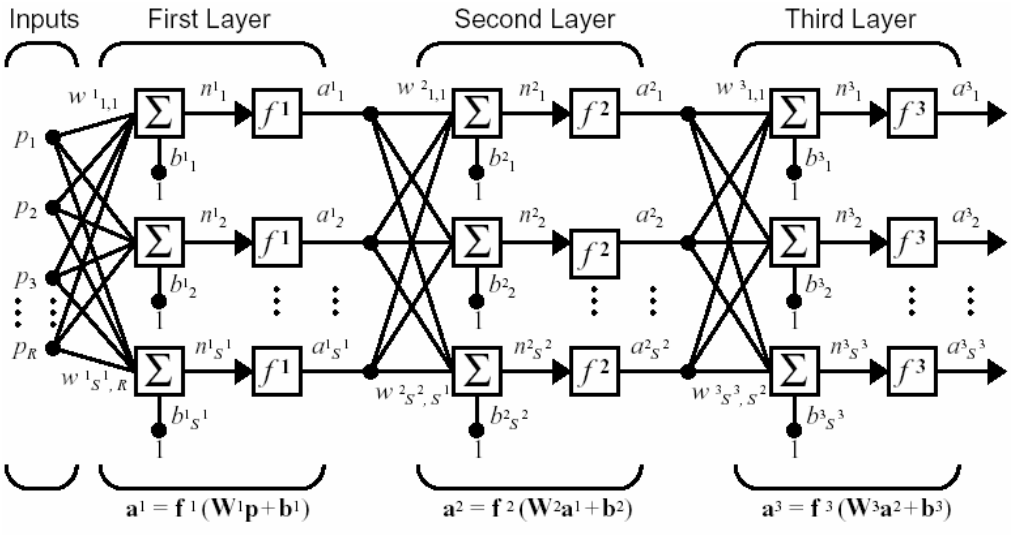


Figure 4.3 Multi Layer Networks [22].

It has R inputs and each layer has a weight matrix W and a bias vector b. The bias and weight matrices corresponding to a particular layer are represented by a superscript above the variable name. The outputs from one layer form the inputs to the next layer. The layer which produces the output is called the *output layer* and the remaining layers are called the *hidden layers*.

The number of inputs to a layer need not be equal to the number of neurons present. The number of neurons and layers are selected depending on the complexity of the problem.

4.4 Feed Forward networks

Feed forward network is a type of network architecture in which the signal travels only in one direction, from input to output. There is no feedback from the output to the input.

4.5 Transfer function

Transfer function can be a linear or a non-linear function of n. The transfer function calculates the desired output a from the net input. In a multilayer network, the transfer function in each layer need not be the same.

4.6 Training the network

Training the network in simple terms, is modifying the weights and biases to perform a particular function. There are many different types of training procedures available. The most commonly used ones are

a. Supervised training.

b. Unsupervised training.

4.6.1 Supervised training

In supervised training, the network is provided with known inputs and the corresponding outputs. When the network is trained, it compares its output with the desired output or target. When the results do not match, the weight and the bias are adjusted so that the network is trained for the desired output.

4.6.2 Unsupervised training

In unsupervised training, the neural network is not provided with the outputs. It trains with the help of inputs. The neural network recognizes a pattern in the input and gets trained for the variation in that pattern.

4.6.3 Training algorithm

The neural network compares its output to the known input. If there is an error between the input and the output, the neural network attempts to minimize the error by adjusting the weights in each layer so that the neural network can generalize well to the given inputs. In other words, the neural network calculates the change in error as each weight is increased or decreased. This process requires the neural network to compute the derivative of weights. In the case of single-layer networks, the error is a linear function of the network weight, hence computing them is easier. In multilayer networks with non-linear transfer functions, computing the errors becomes more difficult.

4.6.4 Performance index

It is very difficult to have an ideal network that can predict the exact output(s) for a given set of input(s). There has to be some measurement tool to identify if sufficient training has occurred and stop training when it has occurred. One such measurement tool is the mean square error (MSE).

In supervised training, the neural network is supplied with known inputs and known targets.

$$\{P_1, t_1\}, \{P_2, t_2\}, \ldots\ldots\ldots \{P_i, t_i\}, \tag{4.5}$$

P_i Is the known input given to the network and t_i is the known target given to the network. As each input is given to the network, it is compared with the output. The mean square error can be represented as

$$f(x) = E(e^2) = E[(t-a)^2]. \tag{4.6}$$

Mean square error is a performance measure that is used by the back-propagation algorithm.

4.6.5 Back propagation algorithm

A multilayer network can be represented as

$$a^{m+1} = f^{m+1}(W^{m+1}a^m + b^{m-1}), \qquad for \qquad m = 0,1,2.....M-1, \tag{4.7}$$

where, M is the number of layers in the network. The input to the network is given by,

$$a^0 = p. \tag{4.8}$$

The outputs of the network is given by

$$a = a^M \tag{4.9}$$

The steepest descent algorithm for the mean square error is given by,

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha \frac{\partial F}{\partial w_{i,j}^m}. \tag{4.10}$$

$$b_i^m(k+1) = b_i^m(k) - \alpha \frac{\partial F}{\partial b_i^m}. \tag{4.11}$$

where, α is the learning rate. For a single-layer network, Eqns (4.10) and (4.11) can be calculated from the following

$$W(k+1) = W(k) + 2\alpha e(k)z(k). \tag{4.12}$$

$$b(k+1) = b(k) + 2\alpha e(k). \tag{4.13}$$

For a multi-layer network, error does not depend upon weights of a single layer but on multiple layers. Chain rule is a solution to this problem.

4.6.5.1 Chain rule

If there is a function f that is an explicit function of n and if the derivative of the function with respect to x is required, then chain rule is used.

An example of the chain rule is described in the following

If $f(n) = e^{2n}$ and $n = 3x^2$ then,

$$\frac{df(n(x))}{dx} = \frac{df(n)}{dn} \times \frac{dn(x)}{dx}, \tag{4.14}$$

$$\frac{df(n(x))}{dx} = 2e^{2n} \times 6x, \tag{4.15}$$

$$\frac{df(n(x))}{dx} = 12xe^{2n}. \tag{4.16}$$

In the same way, using the chain rule, derivative of a function f in Eqns (6.10) and (6.11) can be determined as,

$$\frac{\partial F}{\partial w_{i,j}^m} = \frac{\partial F}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial w_{i,j}^m}. \tag{4.17}$$

$$\frac{\partial F}{\partial b_i^m} = \frac{\partial F}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial b_i^m} \tag{4.18}$$

The function n can be computed as an input to layer m, which can be expressed as a function of weights and biases

$$n_i^m = \sum_{j=1}^{S^{m-1}} w_{i,j}^m \, a_j^{m-1} + b_i^m. \tag{4.19}$$

$$\frac{\partial n_i^m}{\partial w_{i,j}^m} = a_j^{m-1}, \frac{\partial n_i^m}{\partial b_i^m} = 1. \tag{4.20}$$

If we define,

$$s_i^m = \frac{\partial F}{\partial n_i^m}, \tag{4.21}$$

then Eqns (4.17) and (4.18) can be simplified as

$$\frac{\partial F}{\partial w_{i,j}^m} = s_i^m \, a_j^{m-1}. \tag{4.22}$$

$$\frac{\partial F}{\partial b_i^m} = s_i^m. \tag{4.23}$$

Now, the steepest descent algorithm can be expressed as

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha s_i^m a_j^{m-1}. \tag{4.24}$$

$$b_i^m(k+1) = b_i^m(k) - \alpha s_i^m. \tag{4.25}$$

In matrix form Eqns (6.24) and (6.25) become

$$W^m(k+1) = W^m(k) - \alpha s^m (a^{m-1})^T, \tag{4.26}$$

$$b^m(k+1) = b^m(k) - \alpha s^m, \tag{4.27}$$

where,

$$s^m = \frac{\partial F}{\partial n^m} = \begin{bmatrix} \frac{\partial F}{\partial n_1^m} \\ \frac{\partial F}{\partial n_2^m} \\ . \\ . \\ . \\ \frac{\partial F}{\partial n_{s^m}^m} \end{bmatrix}. \tag{4.28}$$

Equation (4.28) gives the sensitivity s^m of just one layer. One has to compute the sensitivities of other layers as well. They are computed from the last layer (output layer) and propagated into the first layer backwards, hence, the term back propagation.

The sensitivities of the other layers can be computed using the Jacobian matrix given by

$$\frac{\partial n^{m+1}}{\partial n^m} = \begin{bmatrix} \frac{\partial n_1^{m+1}}{\partial n_1^m} & \frac{\partial n_1^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_1^{m+1}}{\partial n_{S^m}^m} \\ \frac{\partial n_2^{m+1}}{\partial n_1^m} & \frac{\partial n_2^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_2^{m+1}}{\partial n_{S^m}^m} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \frac{\partial n_{S^{m+1}}^{m+1}}{\partial n_1^m} & \frac{\partial n_{S^{m+1}}^{m+1}}{\partial n_2^m} & \cdots & \frac{\partial n_{S^{m+1}}^{m+1}}{\partial n_{S^m}^m} \end{bmatrix}. \tag{4.29}$$

Therefore, the Jacobian matrix can be written as

$$\begin{aligned} \frac{\partial n_i^{m+1}}{\partial n_j^m} &= \frac{\partial \left(\sum_{l=1}^{S^m} w_{i,l}^{m+1} a_l^m + b_i^{m+1} \right)}{\partial n_j^m} \\ &= w_{i,j}^{m+1} \frac{\partial a_j^m}{\partial n_j^m} = w_{i,j}^{m+1} \frac{\partial f^m(n_j^m)}{\partial n_j^m} = w_{i,j}^{m+1} f(n_j^m), \end{aligned} \tag{4.30}$$

where,

$$f^m(n_j^m) = \frac{\partial f^m(n_j^m)}{\partial n_j^m}. \tag{4.31}$$

The Jacobian matrix can now be expressed as

$$\frac{\partial n^{m+1}}{\partial n^m} = W^{m+1} F^m(n^m). \tag{4.32}$$

where ,

$$F^m(n^m) = \begin{bmatrix} f(n_1^m) & 0 & \dots & 0 \\ 0 & f(n_2^m) & \dots & 0 \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ 0 & 0 & & f(n_{S^m}^m) \end{bmatrix}. \tag{4.33}$$

Using the chain rule again, the recurrence relation can be expressed in matrix form as

$$s^m = F^m(n^m)(W^{m+1})^T s^{m+1}. \tag{4.34}$$

The sensitivities are propagated backwards in the network from the last layer to the first layer. This is expressed as

$$s^M \rightarrow s^{M-1} \rightarrow s^3 \rightarrow s^1. \tag{4.35}$$

The sensitivities at the last layer or output layer is given by

$$s^M = -2F^M(n^M)(t-a). \tag{4.36}$$

4.7 Early Stopping

If the neural network has to precisely predict the underlying function between the inputs and outputs, it should not overfit the given data. Early stopping is a method that is useful to avoid over fitting of data. In this method, the entire input set is divided into training set, testing set, and validation set. The training set is used to compute the weights to obtain the desired output. The validation set is used to determine the time to stop training. As the neural networks starts to overfit the data, the error on the training set starts decreasing

and the error on the validation set starts increasing. The aim of the early stopping algorithm is to minimize the error on the validation set.

CHAPTER 5

PROBLEM STATEMENT

The aim of molecular dynamics (MD) simulations is to model interactions of materials at the atomic scale. For MD simulations of the study of CVD process, it is necessary to critically asses the note of the experimental conditions, precursor gases used, their composition, the reactions involved. This would facilitate in modeling the process.

In a CVD process, the properties of the deposited film depend on its chemical composition. In order to determine the most effective species that aid in the growth of diamond films, MD simulations have been a very common tool, as pointed out in Chapter 2.

A number of precursor atoms and molecules, namely, CH_3, C_2H_2, C_2H_4, CH_2, C, to name a few, may be used as growth species in a CVD process. Diamond film growth occurs by the chemisorption of these radicals on to a diamond surface.

The aims of this investigation are the following

1. To perform MD simulations for investigating the chemisorption, scattering and desorption probabilities of a carbon atom on a diamond (111) surface. We intend to employ the potential energy surface given by Brenner *et al*.

[14] and compute the probabilities of chemisorption, scattering, and desorption for various input parameters.

2. To develop neural networks to reduce the time involved in investigating the effect of input parameters, namely, incidence angle (θ), rotation angle (Φ), impact parameter (b), and kinetic energy (K) of C atom on three events, namely, chemisorption, reflection, and desorption.

3. To use neural networks to predict the underlying function between incidence angle (θ), rotation angle (Φ), impact parameter (b), and kinetic energy of C atom (K) and event probabilities which form the output.

CHAPTER 6

MD SIMULATIONS AND NN IMPLEMENTATION OF EVENT PREDICTION IN A CVD PROCESS

A chemical reaction occurs by the adsorption of a chemical species on a reaction surface. This study focuses on the chemical reactions that occur on a diamond (111) surface. The reactions that are considered in this study are chemisorption, scattering, and desorption. This chapter describes the simulation setup used, the input variables considered, and the output probabilities.

6.1 Model Considered

The present study considers a diamond (111) surface for simulation. The top layer atoms on the diamond surface are covered with hydrogen atoms except for the central one which forms the radical site. The diamond substrate consists of five layers of atoms, out of which the top layer forms the hydrogen layer. Out of the remaining four layers of carbon atoms, the bottom layer consists of boundary atoms for which the Newtonian equations of motion are not integrated. A layer of atoms on each side of the substrate also form boundary atoms. In total, the simulation model consists of 186 atoms, of which 108 atoms are moving atoms. Moving atoms are those for which equations of motion are integrated.

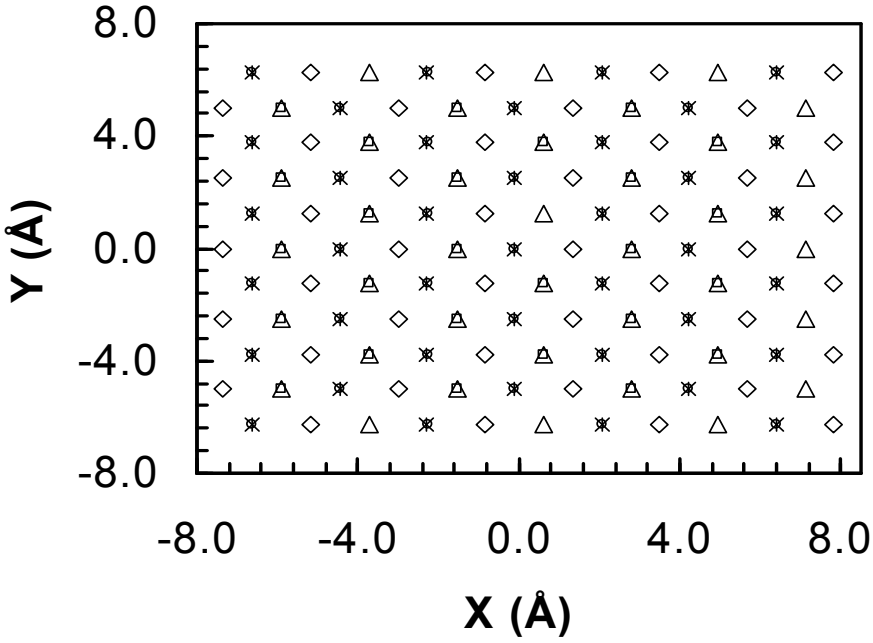


Figure 6.1 Top View of Diamond (111) substrate. In the figure, squares represent top layer hydrogen atoms, triangles represent first layer of carbon atoms, circles represent the second layer of carbon atoms, asterisks represent the third layer carbon atoms, and rhombus represents the fourth layer carbon atoms.

6.2 Input Variables

To study the chemical reactions occurring on the diamond (111) substrate when impacted by a carbon atom, the input variables were selected such that the chemical reaction that occurs simulates the experiment.

The input variables considered are

a. Impact parameter (b),

b. Incident angle (θ),

c. Rotation angle (Φ), and

d. Kinetic energy of carbon atom (K).

Figure 6.2 shows a pictorial view of the input variables considered.

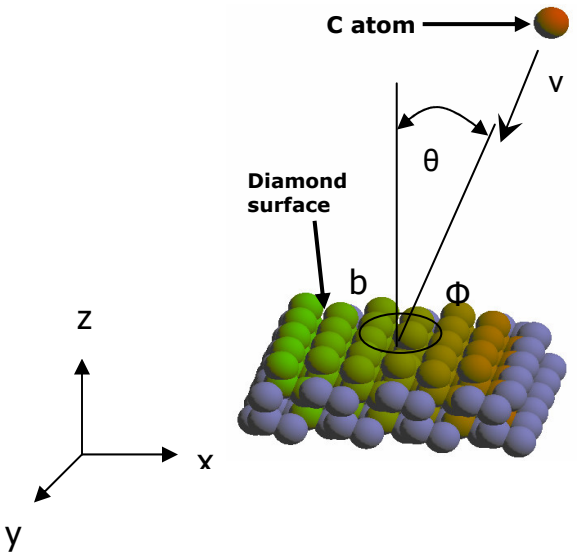


Figure 6.2 Representation of input variables considered

In this study, before the reaction probabilities are studied on the diamond (111) substrate, the substrate has to be brought to the minimum energy configuration. For this, the forces on the atoms in the system are calculated by taking the first derivative of the potential. From the forces, the new positions are obtained by solving the Newtonian equations of motions. If the potential energy of the new structure is greater than the previous one, the velocity of all atoms is set to zero and the process is repeated until the potential energy attains a minimum.

After the substrate is relaxed, kinetic energy equals to 3 kT (where k is the Boltzmann constant and T is the desired temperature), is randomly inserted into each atom and the substrate is allowed to vibrate for ~ 60,000 time steps with a time step of 0.5 fs. The temperature of the substrate is then maintained at 1250 K using a Brendeson thermostat procedure [42] for ~ 50,000 steps.

The procedure for the selection of each of the input variables is given in Sec 6.2.1. In this study, each of these input parameters were held constant for 50 trajectories, which forms one neural network point.

6.2.1 Impact parameter

Impact parameter is the point at which the incident carbon atom is aimed at the substrate. In the present study, the impact parameter is chosen randomly from a distribution function given by [18, 19]

$$b = \sqrt{\zeta_1} \times b_{\max}. \tag{6.1}$$

This is equivalent to the selection from a normalized probability distribution

$$P(b)db = \frac{2bdb}{b_{\max}^2} \tag{6.2}$$

where $b_{\max}$ is the maximum radius around the radical site within which the value of b varies. ζ_1 is a random number that varies between 0 and 1. In this study $b_{\max}$ was found to be 2.5 Å as chemisorption at the radical site was not observed beyond 2.5 Å. The value of b signifies the radius around the radical site where impact can occur. In each trajectory, b was chosen randomly using Eqn 6.1. Figure 6.3 shows a typical distribution of impact parameters generated by Eqn (6.1). Note that the probability of selection varies in a non-linear fashion with b.

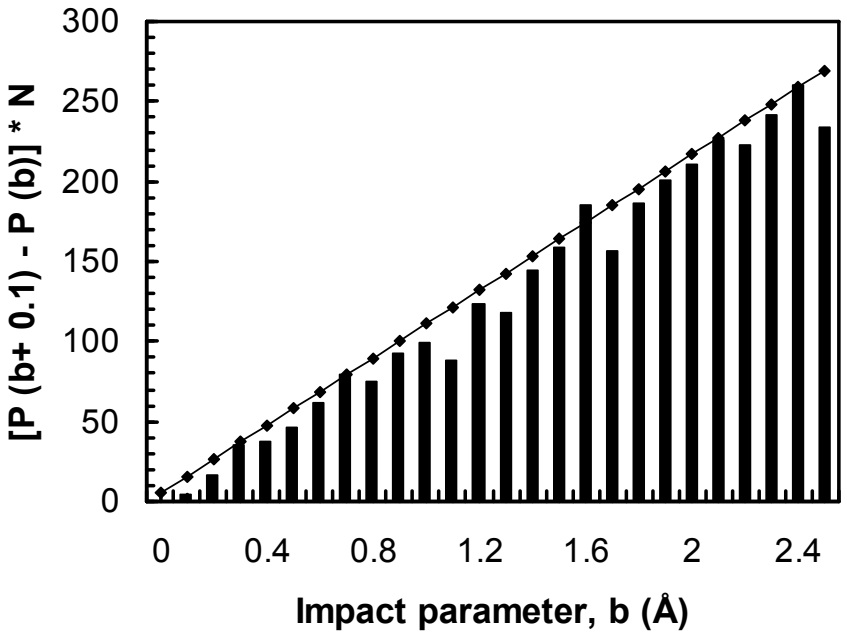


Figure 6.3 Distribution for the impact parameter. Histogram is the plot of distribution. Solid line is the plot of the theoretical function. N represents the number of points (3305 used in this study).

6.2.2 Incident angle

In this study, the incident carbon atom is placed over the substrate at a distance such that there is no influence of the substrate atoms on the incident carbon atom when it is over the edge of the substrate. The incident angle was selected randomly from a distribution weighted by the solid angle element,

$$P(\theta)\, d\theta = Sin\theta\, d\theta \tag{6.3}$$

This selection can be achieved from Eqn (5.4) [18, 19].

$$\theta = \cos^{-1}(1 - \zeta_2(\cos\theta_m), \tag{6.4}$$

where θ_m is the maximum value that the incident angle can take. ζ_2 is a random number which takes values between 0 and 1. θ_m was decided based on the substrate size, and the effect of chemisorption at a distance from the radical site.

In the present study, θ_m was fixed at 21.84°. It was based on the condition that the incident carbon atom should be at least 10 Å above the top layer of the substrate at the start of the trajectory and aimed within 2.5 Å from the radical site, after which there is no effect of chemisorption. Figure 6.4 shows a histogram for incident angle distribution obtained from Eqn (6.4).

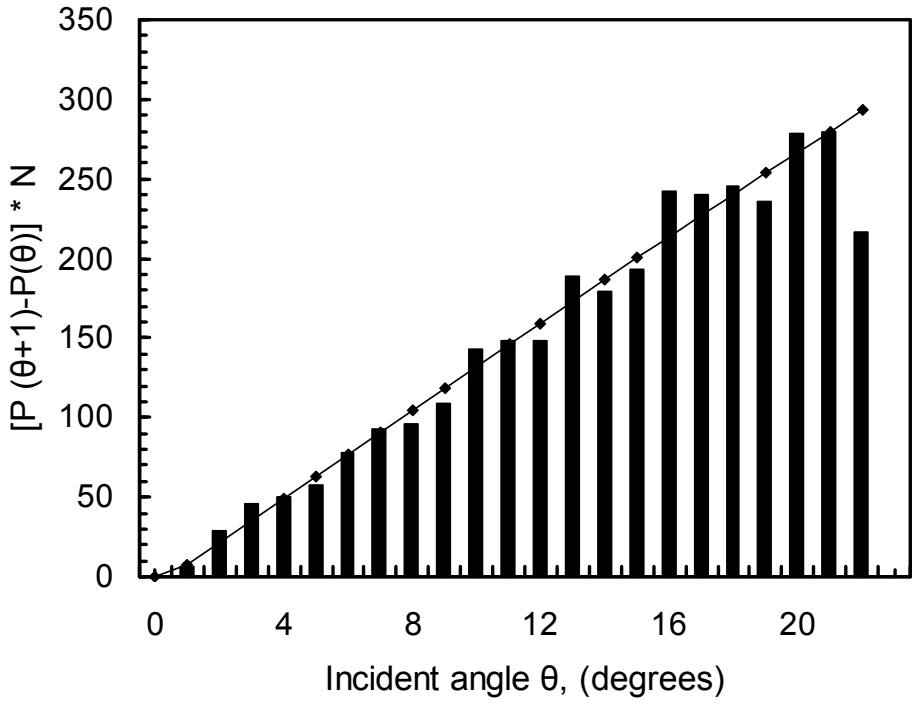


Figure 6.4 Distribution for the incident angle. Histogram is the plot of the distribution obtained. Solid line is the plot of the theoretical function. N represents the number of points (3305 used in this study).

6.2.3 Rotation angle

The rotation angle is defined in Figure 6.2. The rotational angle distribution is uniform over the range $0 \le \phi \le 2\pi$. Therefore, we select its value in each trajectory using Eqn (6.5) [18, 19]

$$\phi = 2 \times \pi \times \zeta_4. \qquad (6.5)$$

ζ_4 is a random number that varies between 0 and 1. Figure 6.5 shows the distribution for the rotation angle ϕ

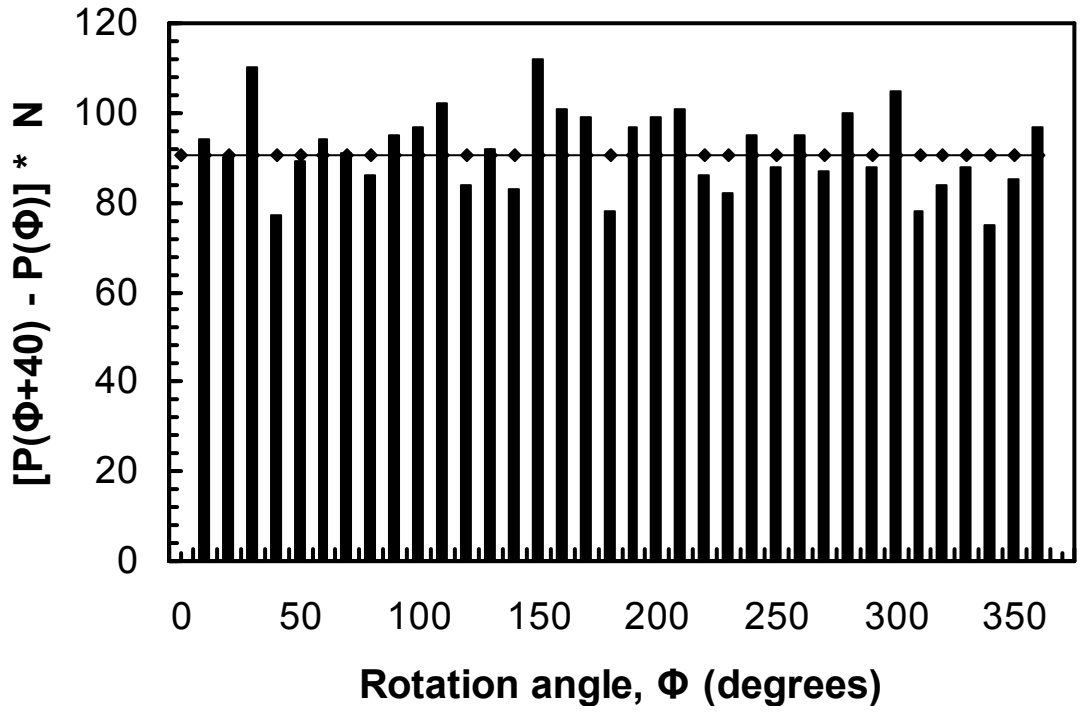


Figure 6.5 Distribution for the rotation angle. Histogram is the plot of the distribution obtained. Solid line is the plot of the theoretical function. N represents the number of points (3305 used in this study).

As can be seen the distribution is essentially uniform.

6.2.4 Kinetic energy of incident carbon atom

Kinetic energy of incident carbon atom (K) is chosen from a Boltzman distribution corresponding to a temperature of 1250 K. The Boltzmann distribution function is given by

$$P(v)dv = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(\frac{-mv^2}{2kT}\right)v^2 dv. \qquad (6.6)$$

Kinetic energy is found using

$$K.E. = \frac{1}{2} m v^2. \tag{6.7}$$

The distribution of Kinetic energy in eV is given in Figure 6.6

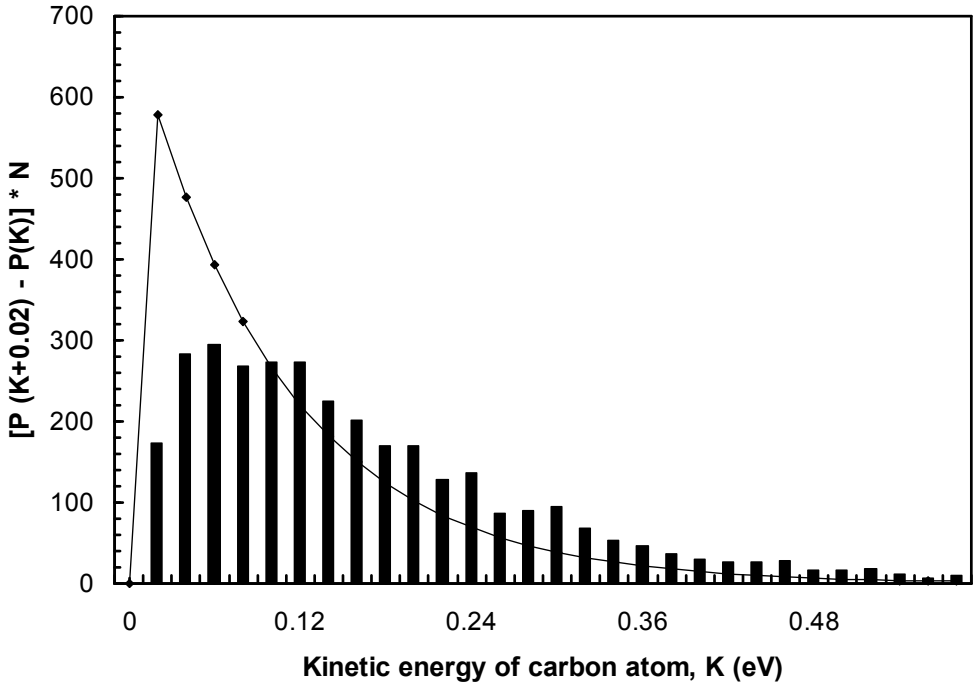


Figure 6.6 Boltzmann distribution for the kinetic energy of incident carbon atom corresponding to 1250 K. Histogram is the plot of the distribution obtained. Solid line is the plot of the theoretical function. N represents the number of points (3305 used in this study).

6.3 Output variables

As mentioned in Section 6.2, 165,250 (3305×50) MD trajectories were run holding all input variables θ, Φ, b, and K constant for a set of 50 trajectories. The event probabilities of interest are

a. Chemisorption probability,

b. Scattering probability, and

c. Desorption probability.

The Z - coordinate and energies associated with each event are described below.

6.3.1 Chemisorption

Chemisorption is bonding between the incident carbon atom and the carbon at the radical site. In the present study, we used the criteria that chemisorption had occurred if the incident carbon atom experienced more than 10 inner turning points and stayed within a bonding distance of 2 Å from the carbon atom at the radical site for at least 1 ps

Figure 6.7 shows the variation of Z - coordinate and potential energy plots with time, as chemisorption occurs. It can be seen from the plot that there is a drop in the potential

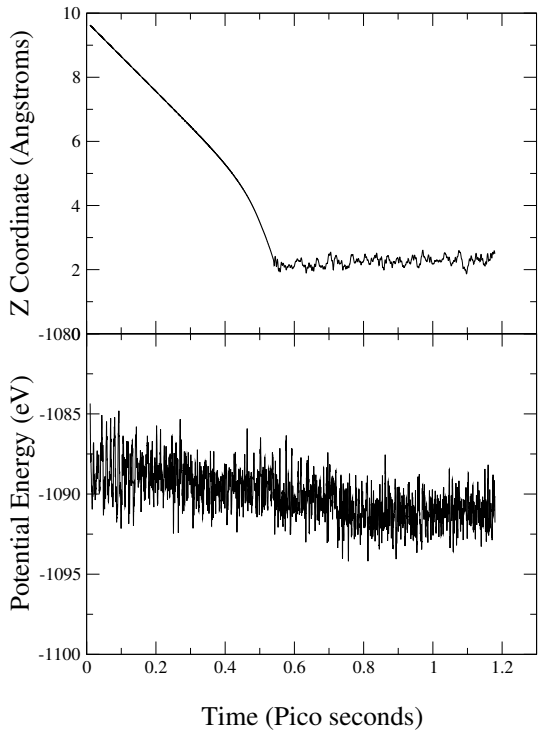


Figure 6.7 Z – coordinate and potential energy plots with time for chemisorption event

6.3.2 Scattering

When the incident carbon atom does not react with the substrate and is reflected, scattering is considered to have occurred. In the present study, if the incident carbon atom exhibits just one inner turning point and is well above the bonding distance of 2 Å at the end of 1 ps after it has reached the substrate, scattering is considered to have occurred.

Figure 6.8 shows the Z - coordinate plot and the potential energy plot with time, as scattering occurs. It can be seen from the plot that there is no drop in potential energy

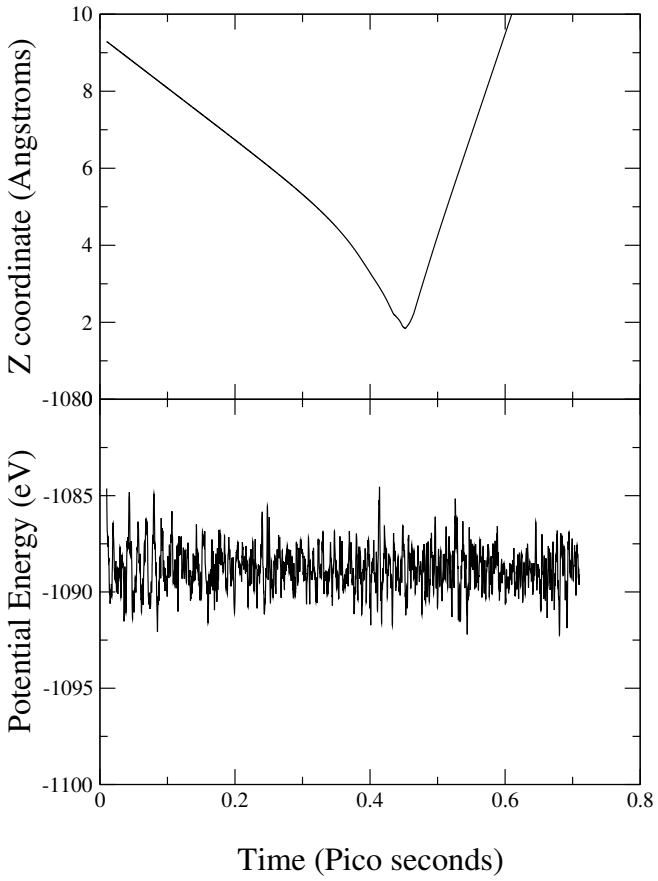


Figure 6.8 Z – coordinate and potential energy plots with time for scattering event

6.3.3 Desorption

When the incident carbon atom does not react with the substrate but stays within the bonding distance for sometime and then scatters away, desorption is assumed to have occurred. In the present study, if the incident carbon atom exhibits 2 to 9 inner turning points and is well above the bonding distance of 2Å from the radical site at the end of 1 ps after it has reached the substrate, desorption is assumed to have occurred.

Figure 6.9 shows the Z - coordinate plot and potential energy plot with time as desorption occurs. No significant drop in potential energy is seen as there is no reaction oc

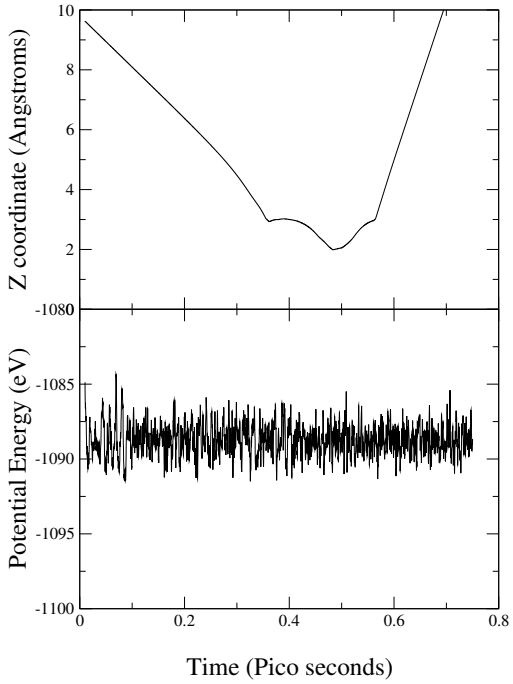


Figure 6.9 Z – coordinate and potential energy plots with time for desorption event

Average chemisorption, scattering, and desorption probabilities were alculated from the results of 50 trajectories with the inputs θ, Φ, b, and remaining constant.

The probability from 50 MD trajectories for a given input (θ, Φ, b, and K) corresponds to one set of output and input to the neural network. 3305 sets of 50 trajectories were run and the corresponding inputs and outputs were used for training the neural network.

6.4 Implementation of Neural Networks

The main purpose for the use of neural networks in this study is that artificial neural networks operate in parallel and perform a multitude of activities at the same time. As discussed in Chapter 2, MD simulations are computationally intensive. Also the study of probabilities involves statistical errors. Hence, use of neural networks for this study is well justified as time for the computation and statistical errors are immensely reduced.

In this study, there are three networks used to predict the probabilities for,

a. Chemisorption

b. Scattering

c. Desorption

θ, Φ, b, and K form the inputs to each of the neural networks and event probabilities form the output of the network. The feed forward multilayer neural networks used consist of one input layer and one hidden layer. Fifty neurons are used in the hidden layer.

The probabilities supplied to the neural network as targets are from 50 MD trajectories for the same input conditions. The 50 trajectories provide an average over the vibrational phase of the lattice.The output of the neural network is the average of 50 neural networks.

6.4.1 Transfer functions used

In the first layer or the hidden layer, the tan sigmoid transfer function as shown in Figure 6.10 was used. This transfer function takes values between -∞ and +∞ and outputs values between -1 and +1

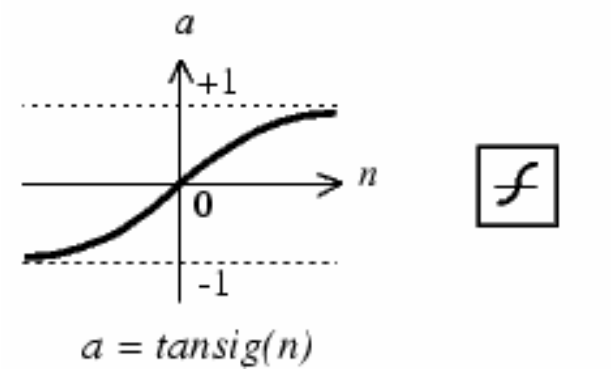


Figure 6.10 Tan sigmoid transfer function [22].

The tan sigmoid function takes the form,

$$a = \frac{2}{(1+e^{-2n})} - 1. \tag{6.8}$$

In the output layer purelin transfer function was used. This is a linear transfer function and the output of the transfer function is equal to its input. Figure 6.11 is a representation of the purelin transfer function [22].

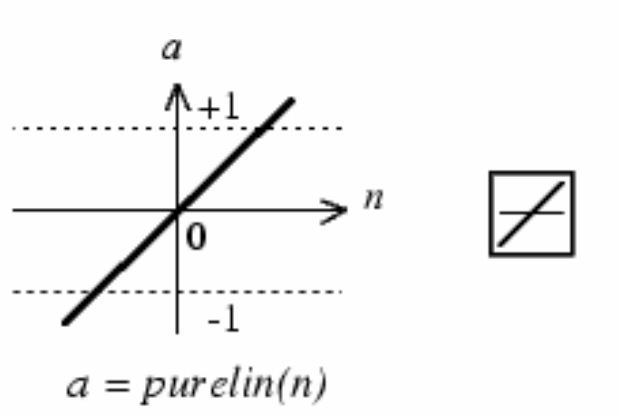


Figure 6.11 Linear transfer function [22].

The purelin transfer function takes the form,

$$a = n. \tag{6.9}$$

6.5 Normalization

The transfer function tan sigmoid used in the hidden layer takes up only values between -1 and 1 and gives outputs between -1 and 1. Hence, the values that are fed into the network have to be scaled between -1 and +1. The algorithm used for normalization is given by

$$p_i = 2 \cdot \frac{(p_i - p_{\min})}{(p_{\max} - p_{\min})} - 1, \tag{6.10}$$

where, $p_{\min}$ and $p_{\max}$ are the minimum and maximum of the input and target values, respectively. For the present study, a total of 3305 points have been used to train the neural network, of which 85% (2809 points) are chosen for training and 15% (496 points) are chosen for validation. The variables which are to be studied forms the testing set.

CHAPTER 7

RESULTS AND DISCUSSION

As stated in Chapter 6, MD simulations were carried out on a diamond (111) substrate and the reaction probabilities were determined. These probabilities were obtained by running a set of 50 trajectories to average over lattice vibrational phases. The inputs and outputs (reaction probability) corresponding to 50 trajectories formed one neural network training or validation point. In this study, 3305 points were generated for the neural network. The neural network was then trained using these points. The inputs to the neural network were θ, Φ, b, and K and the outputs were

a. Chemisorption probability,

b. Scattering probability, and

c. Desorption probability.

The main aim of this study is to reduce the time involved in studying the effect of the complete range of input parameters on the output by using neural networks trained on MD data.

Three neural networks were developed to study each of the event probabilities.

7.1 Training and testing plots

The neural network was trained using the early stopping algorithm as discussed in Chapter 6. Figures 7.1 to 7.3 show the training and testing plots for

chemisorption, scattering, and desorption probabilities, respectively. In all these plots, the X - axis represents the probability predicted by MD , while the Y – axis represen

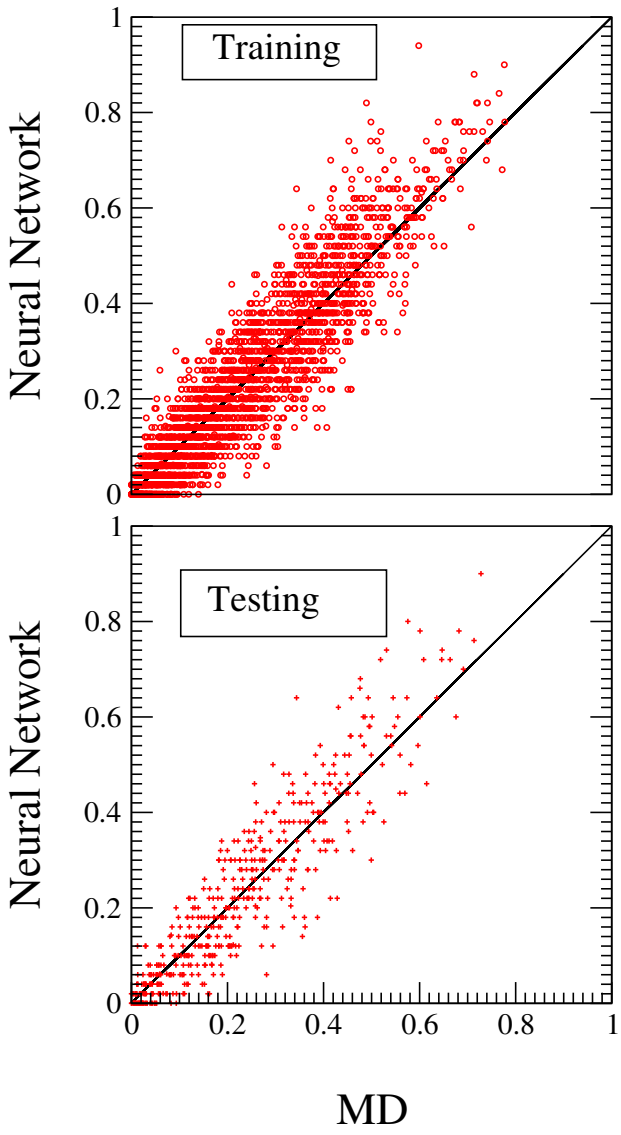


Figure 7.1 Training and testing plots for chemisorption probability

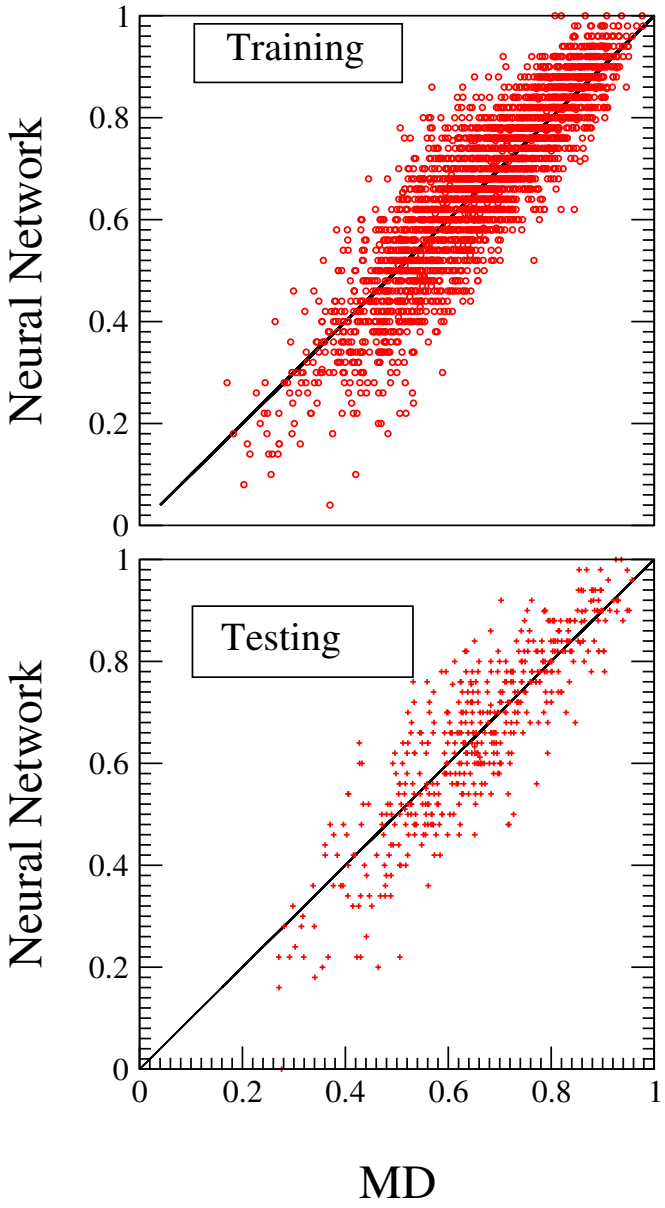


Figure 7.2 Training and testing plots for scattering probability

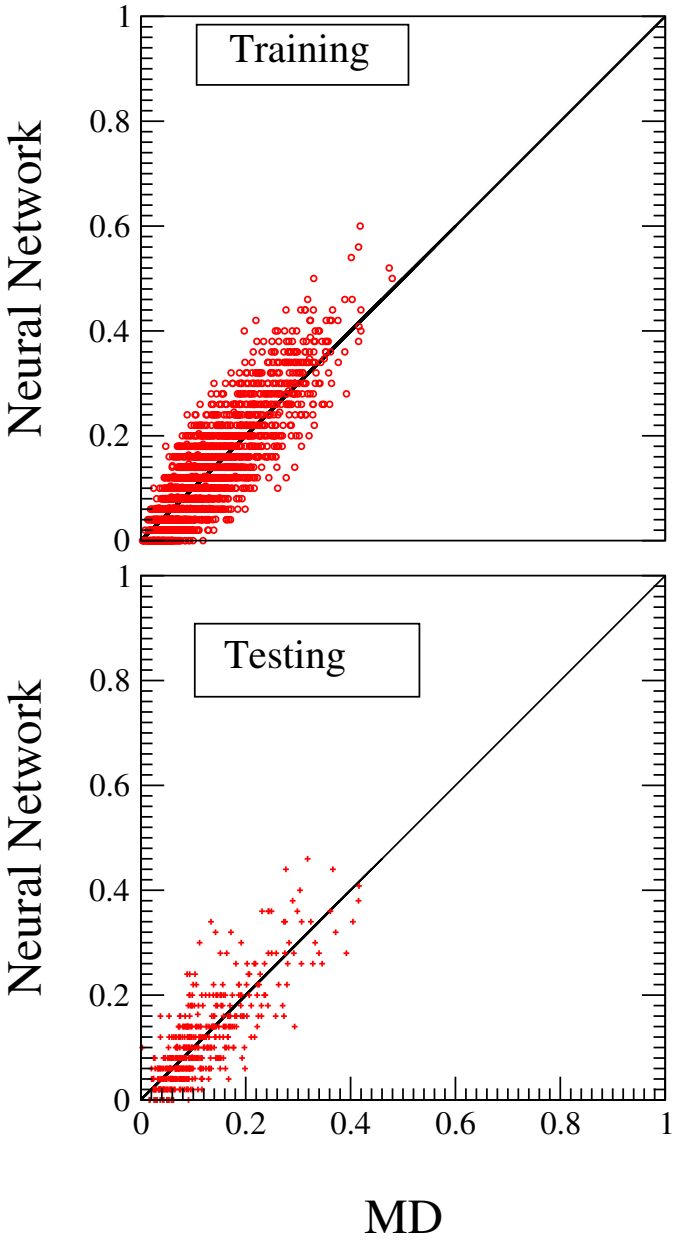


Figure 7.3 Training and testing plots for desorption probability

Out of 3305 points, 70% were used for training the network, another 15% were used to test the network, and the remaining 15% were used for validation. In Figures 7.1 to 7.3, the Y - axis represents the output of the neural network while the X - axis represents the MD data. If the neural network predicts the MD data exactly, the plot would fit the 45 ° line.

7.2 Neural network predictions and the results of molecular dynamics (MD)

After the network had been trained, it was tested with the already available MD data. The MD data were generated by varying one of the input parameters while holding the others constant. The statistical error for one sigma standard deviation on MD data is calculated using Eqns (7.1) to (7.2) [18]

$$\%error = \sqrt{\frac{N_T - N_s}{N_T * N_s}} \times 100, \tag{7.1}$$

where, N_T is the total number of events (50 in this study) and N_s is the total number of successful events

For one sigma limit, the error is given by

$$\sigma_e = P_{event} \pm error. \tag{7.2}$$

By this method, one sigma error on each molecular dynamics data was calculated. There is a 67% chance for neural network predictions to lie within one sigma standard deviation limit of MD. The agreement of neural network predictions with MD is described in this section.

For better accuracy in the neural network results, while training the network, out of the 3305 points, 85% were used for training and 15 % were used for validation. The additional MD points generated formed the testing set for the network. To reduce random statistical error, 50 neural networks were trained. Each training represented one neural network. For each neural network, the initial weight matrices were randomly chosen. Also, the 85% data for the training set and remaining 15% data as a validation set were chosen randomly for each network from the 3305 points. This type of random selection of weights and

training points were made to allow the 50 neural networks to average out the random statistical error in the data. For all of the 50 neural networks, the additional MD points generated formed the testing set. The output of the neural network has been obtained by averaging over all 50 neural networks.

Figures 7.4 to 7.7 show the ability of the neural network to predict the testing data, which were MD points generated in addition to 3305 points. As mentioned above, the neural network results have been obtained by averaging 50 neural networks.

Figure 7.4 shows the effect of impact parameter (b) on the probabilities under study. In this case, the impact parameter (b) is varied from 0 to 2.5 Å while the other constant input parameters are incident angle (θ) = 15°, rotation angle (Φ) = 200°and kinetic energy of carbon atom (K) = 0.04 eV. It can be seen from Figure 7.4, that as the point of impact of the incident carbon moves away from the radical site, it is difficult to form a bond at the radical site. Therefore, the probability of chemisorption P_c decreases, whereas scattering probability P_s increases. The probability of desorption P_D is less near the radical site as probability of chemisorption is higher, whereas, it increases and remains constant with increase in distance from radical site.

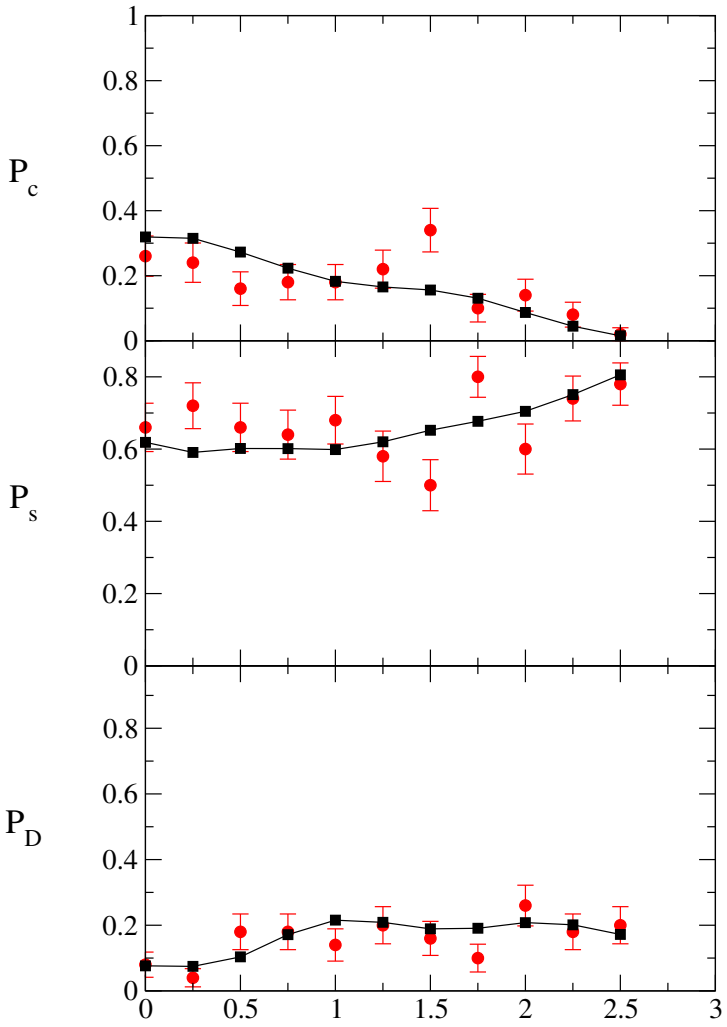


Figure 7.4 Neural Network and MD predictions of effect of the impact parameter (b). P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively. Circles and squares denote MD and neural network data, respectively. The error bars represent one sigma limit of statistical uncertainty.

Figure 7.5 shows the effect of incident angle (θ) on the probabilities under study. In this case, the incident angle is varied from 1° to 21°. The other constant input parameters are impact parameter (b) = 1.5 Å, rotation angle (Φ) = 200° and kinetic energy of carbon atom (K) = 0.14 eV. As the angle of incidence of the

carbon atom increases, chemisorption probability P_c decreases as the incoming carbon atom sees a hydrogen atom in its way. A similar phenomenon for chemisorption probability on diamond (111) surface was observed by Hu *et al*. [21, 41]. Also, with increasing incident angle (θ), scattering probability P_s increases as there is a greater probability for the incoming carbon atom to get scattered away by the surface hydrogen. The desorption probability P_D decreases with increase in incident angle, as there is a high probability of not forming a bond at the radical site.

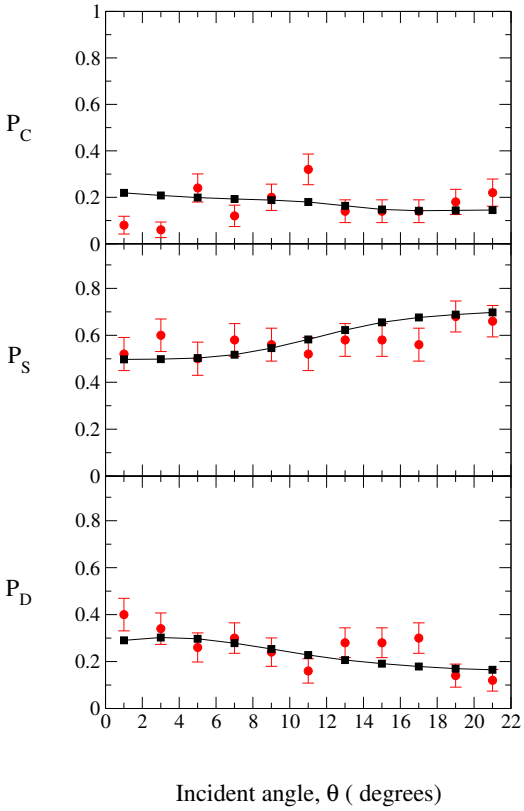


Figure 7.5 Neural Network and MD predictions of effect of the incident angle (θ).$P_{C,}$ P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively. Circles and squares denote MD and neural network data, respectively. The error bars represent one sigma limit of statistical uncertainty.

Figure 7.6 shows the effect of kinetic energy of the incident carbon atom (K) on the probabilities $P_{C,}$ P_S and P_D. In this case, the kinetic energy (K) was varied from 0.005 to 0.8 eV. The other constant input parameters were impact parameter = 1.5 Å, Rotation angle (Φ) = 200°, and incident angle (θ) = 15°. It can be seen that the chemisorption probability $P_{C,}$ increases with increase in kinetic energy. Less chemisorption probability at lower kinetic energy of carbon atom is due to a barrier that hinders chemisorption. Also, a trend in the chemisorption P_C probability and the scattering probability Ps could be the seen from Fig 7.6. The increasing trend in chemisorption probability P_C initially can be attributed to the fact that, at lower kinetic energies, the carbon atom might get repelled by the surface hydrogen atoms and at higher energies, the incident carbon atom might have higher energy to overcome repulsion by hydrogen atoms. Following the increasing trend in chemisorption probability $P_{C,}$ there is a slight decrease. This could be due to the fact that at relatively higher kinetic energies, the incoming carbon atom is unable to dissipate its energy into the lattice and hence a low chemisorption probability. The decreasing trend is again followed by an increasing trend. At higher kinetic energies of the carbon atom, there is partial C - H bond formation between the incoming carbon atom and the hydrogen atom present on the surface of the substrate. This partial bond formation facilitates energy transfer and thereby there is a higher possibility of carbon atom chemisorbing due to lesser energy. A complimentary trend could be observed in the scattering probability P_S. As for the desorption probability P_D, there is an effect of hydrogen repulsion and radical site attraction. Therefore, the desorption

probability P_D, is relatively high at lower kinetic energies and it decreases with higher kinetic energies as chemisorption probability P_C increases. A similar trend of overall increase in chemisorption probability with increase in kinetic energy was observed by Alfonso *et al*.[7], Du *et al*.[13], Gernster *et al*.[8], Huang *et al*. [10] and Neyts *et al.* [9] by studies on various diamond surfaces using different incident hydrocarbon species.

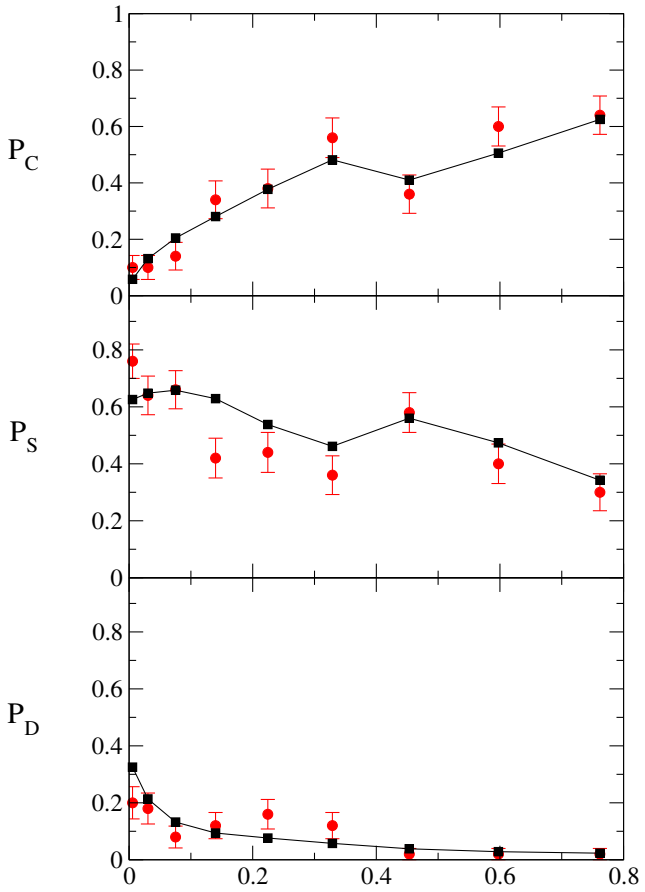


Figure 7.6 Neural Network and MD predictions of effect of the kinetic energy (K) of incident carbon atom. P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively. Circles and squares denote MD and neural network data, respectively. The error bars represent one sigma limit of statistical uncertainty.

Figure 7.7 shows the effect of rotation angle (Φ) on the probabilities P_C, P_S and P_D. In this case, rotation angle (Φ) was varied from 0 to 360°. The other constant input parameters were impact parameter 2 Å, incident angle (θ) = 15°and kinetic energy of carbon atom (K) = 0.14 eV. There seems to be no effect of approach angle on the reaction probabilities.

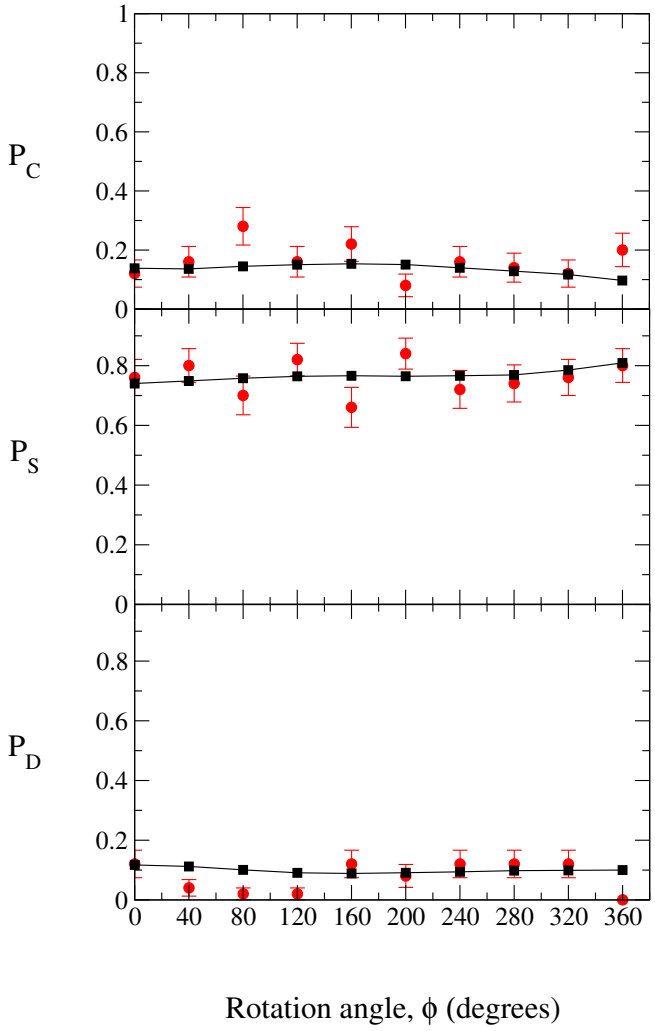


Figure 7.7 Neural Network and MD predictions of effect of the rotation angle (Φ). P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively. Circles and squares denote MD neural network data, respectively. The error bars represent one sigma limit of statistical uncertainty.

It could be inferred from Figures 7.4 to 7.7 that the predictions of the trained neural network are in agreement with the results given by MD.

The neural network predictions shown in Figures 7.4 to 7.7 are the average of 50 neural networks. Therefore, the neural network predictions also have a variation in them. Figure 7.8 is the variation of P_c with impact parameter in neural network and MD.

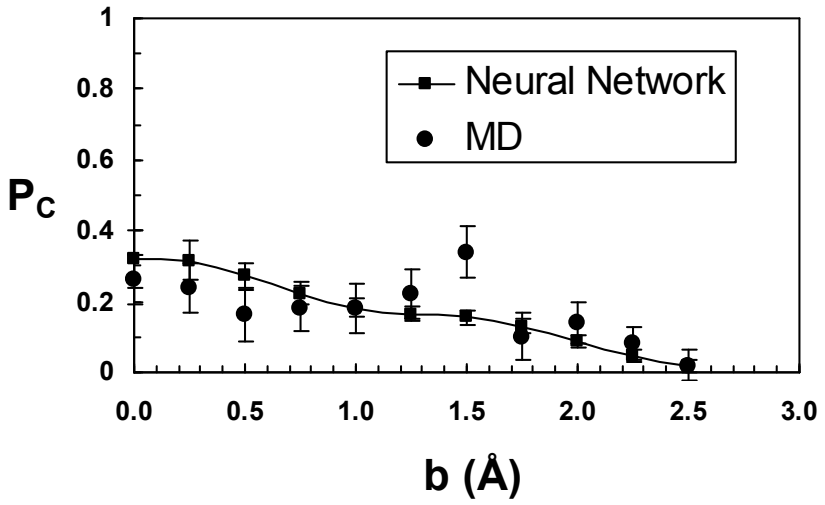


Figure 7.8 Statistical fluctuations in MD and Neural network

The error bars corresponding to the circles and the error bars corresponding to squares indicate the error of MD data and neural network data, respectively.

The following can also be observed from Figures 7.4 to 7.8

1. The neural network is able to follow the trend predicted by MD convincingly.

2. The prediction of neural network is a smooth curve with less statistical fluctuations than MD.

7.3 Neural Network predictions of event probabilities

From Figures 7.4 to 7.7, it is evident that the neural network can predict the underlying function between inputs and outputs. Therefore, the neural network trained on MD data can be used to explore a wide input range at the expense of very small CPU time. Figures 7.9 to 7.12 show such results. Unless otherwise specified, the results described in Figures 7.9 to 7.12 correspond to the following values of input parameters.

a. Rotation angle (Φ) = 200°

b. Incident angle (θ) = 15°

c. Impact parameter (b) = 1.5 Å

d. Kinetic energy of carbon atom (K) = 0.14 eV

From Figure 7.9 and 7.10, it could be inferred that the rotation angle (Φ), has no effect on the reaction probabilities $P_{C,}$ P_S and P_D.

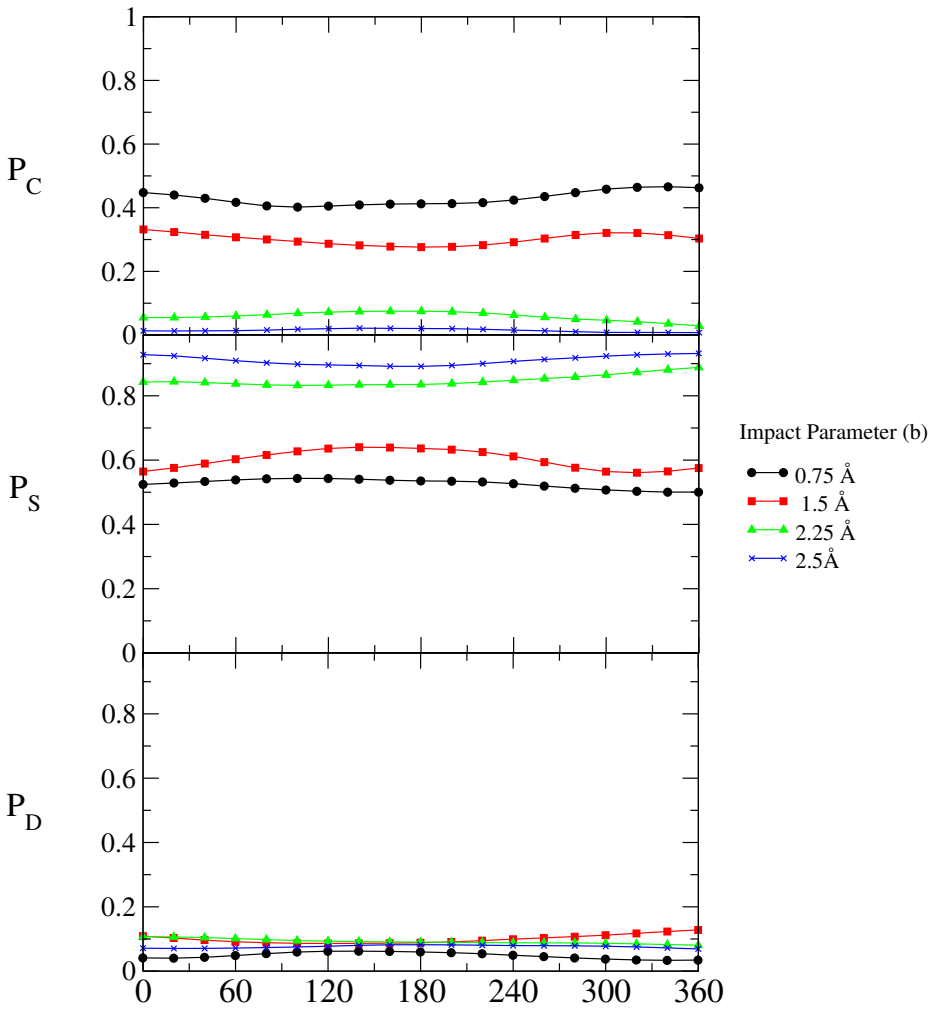


Figure 7.9 Neural Network plot for the study of effect of rotation angle (Φ) on probabilities. P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively.

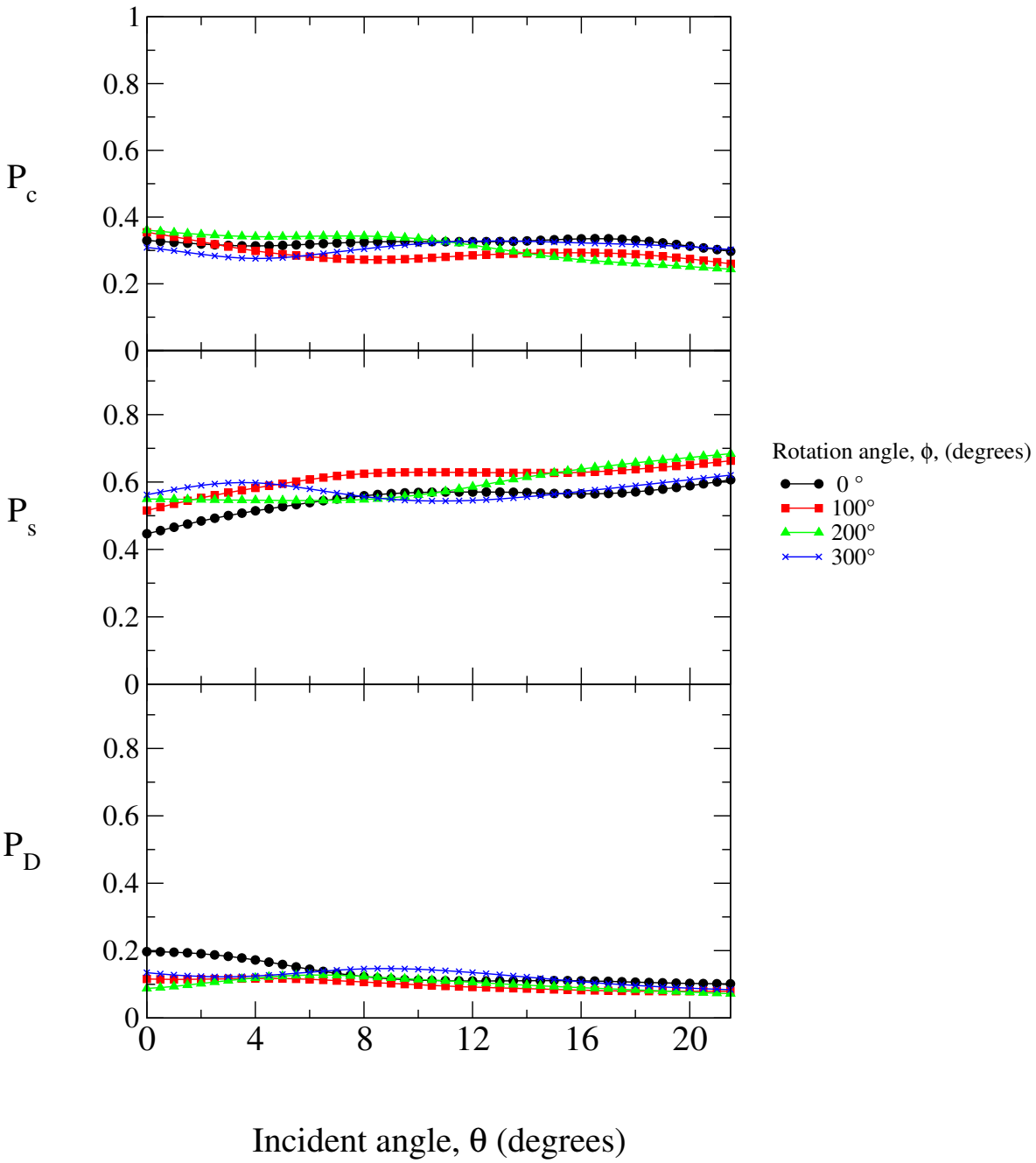


Figure 7.10 Neural Network plot for the study of effect of incident angle (θ) on probabilities. P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively.

From Figure 7.11, it can be seen that at impact parameter values close to the radical site, the chemisorption probability P_c is less at lower incident angles and it is high with higher angle of incidence and a complimentary trend could be observed in the scattering probability P_S . This may be due to the fact that, as the incident carbon atom strikes the radical site head on or at very low incident angles (of the order of 5-10°, in this case), there is not much time for the incident carbon atom coming in at a higher velocity to form a bond at the radical site .The kinetic energy of the carbon atom is higher than the bond energy of C-C bond and therefore no chemisorption occurs. In the case of higher incident angles, the incident carbon atom loses much of its energy on to the vibrating hydrogen atoms before it strikes the radical site and hence a higher chemisorption probability P_C could be attributed to higher incident angles. Figure 7.12 shows the neural network predictions of the effect of kinetic enrgy and the incident angle of the incoming carbon atom on the probabilities. P_C, P_S and P_D are the probabilities of chemisorption, scattering, and desorption respectively. The explanation given to Figure 7.6 can be extended to each curve in Figure 7.12 as well.

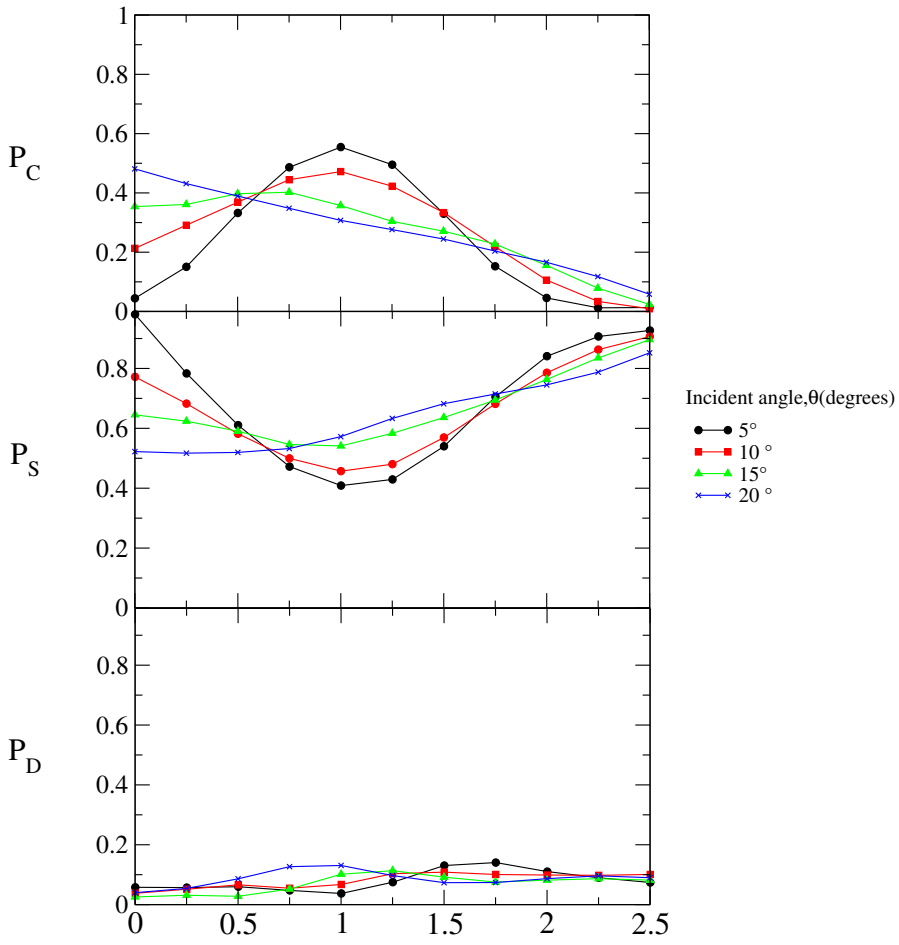


Figure 7.11 Neural Network plot for the study of effect of impact parameter (b) on probabilities. P_C, P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively.

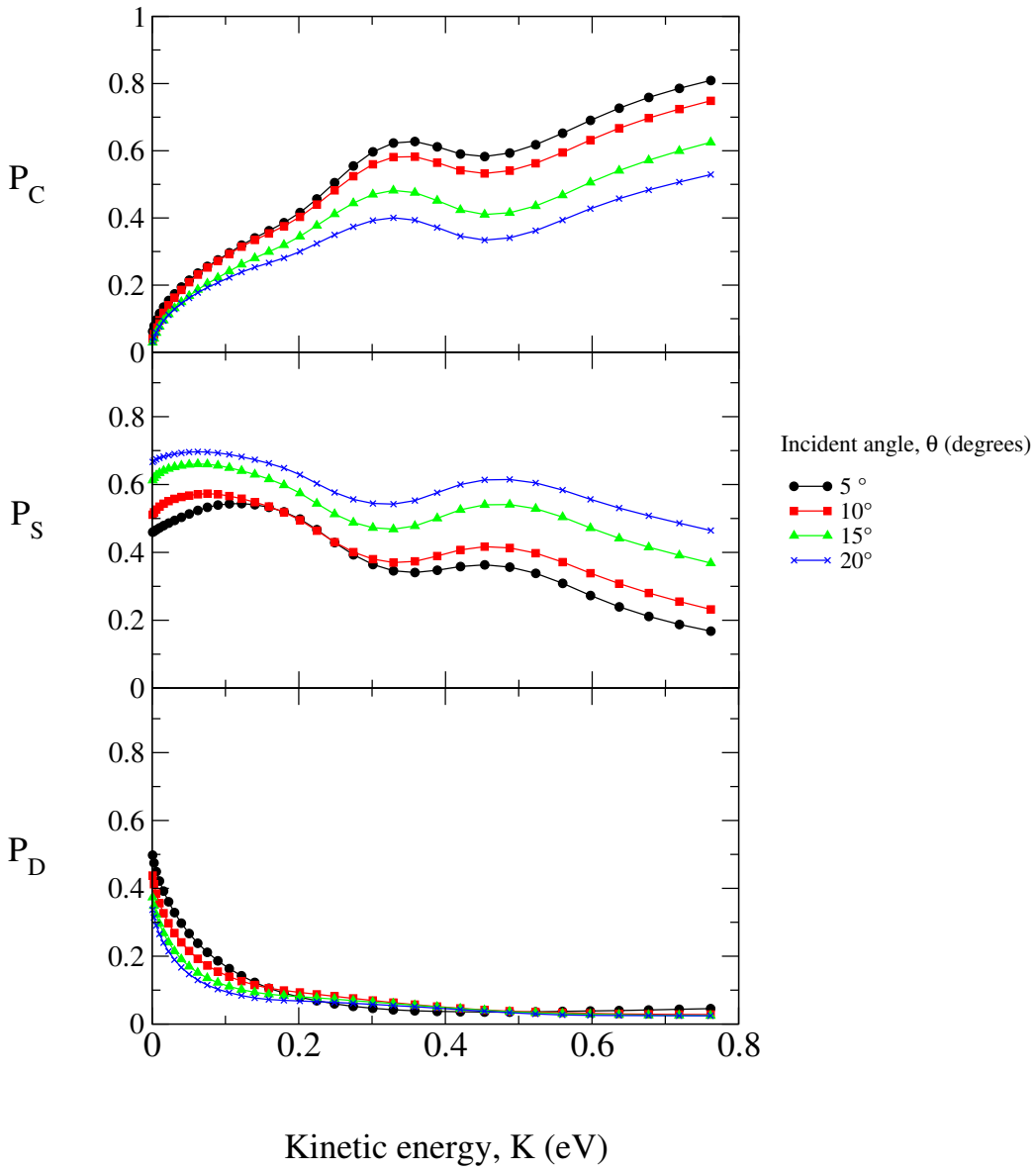


Figure 7.12 Neural Network plot for the study of effect of kinetic energy of the incident carbon atom (K) on probabilities. $P_{C,}$ P_S and P_D, are the probabilities of chemisorption, scattering, and probability of desorption, respectively.

7.4 A comparison of CPU time required

Neural Network is found to be very efficient in terms of processing time compared to MD. In this study, it was found that it takes approximately 20 minutes of CPU time to generate a single MD point. This MD point contains

information of all three probabilities. The time taken for neural network prediction is essentially unaffected by the number of points. It takes less than an hour of CPU time for the neural network to train and predict the probabilities for chemisorption, scattering, and desorption. Depending on the number of points to be simulated, the time taken for the neural network to predict these values slightly increases, but is insignificant compared to the time taken to run MD. Although, clock time to run MD trajectories depends on computer processing speed, it can be said that, with the same computer processing speed, neural network would take less time than MD (minutes compared to days). This can be noted from Figure 7.13

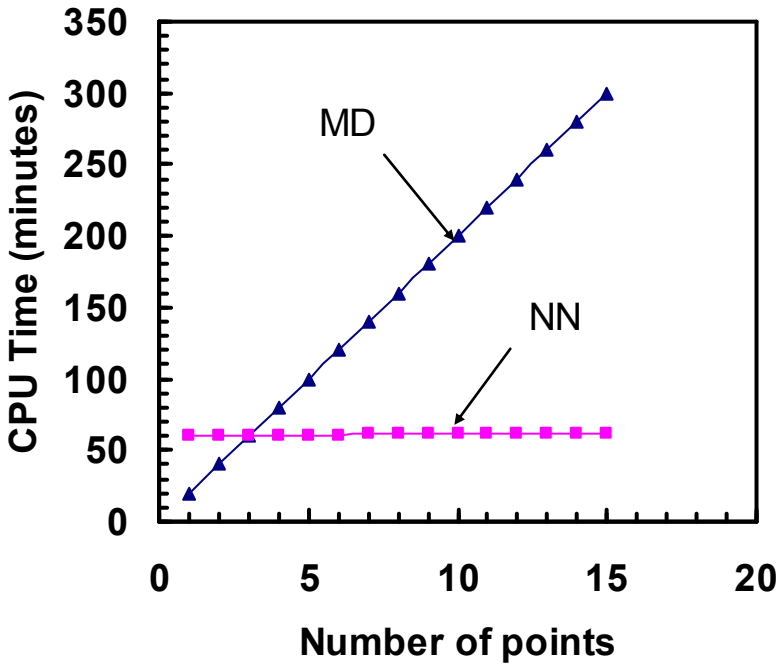


Figure 7.13 CPU time requirements – MD and Neural network

CHAPTER - 8

CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

Chemical vapor deposition of diamond produces precise and uniform coatings. As the coatings are of the order of nanometers, the mechanism for formation of such a coating cannot be studied experimentally. Simulation tools have to be used for such studies. Molecular dynamics is one such simulation tool that has been used over years to study the reaction probabilities of radicals CH_3, C_2H_2, and C_2H_4 on diamond surfaces [17-19]. However, it has a drawback in that such simulations require enormous amount of CPU time.

In this study, an effort was made to reduce the aforesaid drawback of molecular dynamics by using neural networks. Molecular dynamics (MD) simulations were performed to study the chemisorption, scattering, and desorption probabilities of a carbon atom on a diamond (111) surface. Results obtained from MD simulations were then used to train a multilayer neural network. The conclusions that may be drawn from this study are given in the following

1. The trained neural network is able to predict the probabilities at a much faster rate than MD

2. A trained neural network can easily be used to study the effect of any input parameter over a wide range of other parameters which is not a simple task with MD.

3. The predictions of the neural network are reasonably accurate.

4. As impact parameter increases chemisorption probability (P_c) decreases whereas the scattering probability (P_s) increases and desorption probability (P_D) decreases.

5. As incident angle (θ) increases, chemisorption probability (P_c) decreases, whereas scattering probability (P_s) increases and desorption probability (P_D) decreases.

6. As kinetic energy of the carbon atom (K) increases, chemisorption probability (P_c) increases, whereas scattering probability (P_s) decreases and desorption probability (P_D) decreases.

7. As rotational angle (Φ) increases, there seems to be no effect on any of the three probabilities, namely, chemisorption probability (P_c), scattering probability (P_s), and desorption probability (P_D).

8.2 Future work

Having established that the neural network can be used to study the effect of parameters faster than MD, a large domain remains to be explored. The scope of this study can be further extended to the following

1. Input variables for this study had been selected from distribution functions to simulate the physical experiment. The study of a uniform

distribution of the input parameters may also be an interesting domain to explore.

2. Temperature of the substrate can be varied and its effect on the probabilities can be studied.

3. Reaction probabilities of different species can be studied.

4. Reaction probabilities on surfaces without capped hydrogen atoms can be studied.

REFERENCES

1. Choy, K. L., “Chemical vapor deposition of coatings,” Progress in Materials Science **48** (2003) 57-170.
2. Kulisch, W., Popov, C., Boycheva, S., Buforn, L., Favaro, G., and N.Conte, “Mechanical Properties of nanocrystalline diamond/ amorphous carbon composite films prepared by microwave plasma chemical vapor deposition,” Diamond and Related Materials **13** (2004) 1997-2002.
3. Lindstam, M., Wanstrand, O., Boman, M., and K. Piglmayer, “Mechanical and Tribological aspects on a- C films deposited by lamp assisted chemical vapor deposition,” Surface and Coatings Technology **138** (2001) 264-268.
4. Zhang, Z. M., Shen, H. S., Sun, F. H., He, X. C., and Y. Z. Wan, “Fabrication and application of chemical vapor deposition diamond coated drawing dies,” Diamond and Related Materials **10** (2001) 33-38.
5. Ahmed, W., Sein, H., Jackson, M., and R. Polini, “Chemical vapor deposition of diamond films onto tungsten carbide dental burs,” Tribology International **37** (2004) 957-964.
6. Gäbler, J., Schäfer, L., and H. Westermann, “Chemical vapor deposition diamond coated microtools for grinding, milling and drilling,” Diamond and Related Materials **9** (2000) 921-924.

7. Alfonso, D. R., and S. E. Ulloa, “Molecular- dynamics simulations of methyl – radical deposition on diamond (100) surfaces,” Phys. Rev. B **48** (1993) 12235-12239.
8. Gerstner, E. G., and B. A. Pailthorpe, “Molecular dynamics simulation of thin film amorphous carbon growth,” Journal of Non-crystalline solids **189** (1995) 258-264.
9. Neyts, E., Bogaerts, A., Gijbels, R., Benedikt, J., and M. C. M van de Saden, “Molecular dynamics simulations for the growth of diamond-like carbon films from low kinetic energy species,” Diamond and Related Materials,” **13** (2004) 1873-1881.
10. Huang, Z., Pan. Z. Y., Wang, Y. X., and A. J. Du, “Deposition of hydrocarbon molecules on diamond (001) surfaces: atomic scale modeling,” Surface and Coatings Technology 158-159 (2002) 94-98.
11. Alfonso, D. R., Yang, S. H., and D. A. Drabold, “*Ab initio* studies of hydrocarbon adsorption on stepped diamond surfaces,” Phys.Rev.B **50** (1994) 15369 – 15380.
12. Alfonso, D. R., and S. E. Ulloa., “Molecular – dynamics simulations of methyl – radical deposition on diamond (100) surfaces,” Phys.Rev.B **48**(1993) 12235 – 12239.
13. Du. A. J., Pan, Z. Y., Ho, Y. K., Wang, Y. X., and Y. Xu, “ Energy window effect in chemisorption of C_{36} on diamond (001) – (2×1) surface,” Nuclear Instruments and methods in Physical Research B **180** (2001) 153 -158.

14. Brenner, D. W., Shenderova, O. A., Harrison, J. A., Stuart, S. J., Ni, B., Susan B Sinott, " A Second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons ", J. Phys.: Condens. Matter **14** (2002) 783-802.
15. Huang, Z., Pan, Z. Y., Zhu, W. J., Wang, Y. X., and A. J. Du, "Energy Dependence of methyl – radical adsorption on diamond (001) – (2 $\times$1) surface," Surface and Coatings Technology **141** (2001) 246-251.
16. Kaukonen, M. O., and R. M. Nieminen, " Molecular – dynamics simulation of the growth of diamond – like films," Surface Science **331 – 333** (1995) 975 – 977.
17. Neyts, E., Bogaerts, A., Gijbels, R., Benedikta, J., and M.C.M van de Sanden, "Molecular Dynamics simulation of the impact behaviour of various hydrocarbon species on DLC," Nuclear Instruments and Methods in Physics Research B **228** (2005) 315 – 318.
18. Perry, M. D., and L. M. Raff, "Theoretical Studies of Elementary Chemisorption Reactions on an Activated Diamond Ledge Surface," J.Phys.Chem **98** (1994) 4375-4381.
19. Perry, M. D., and L. M. Raff, "Theorotical Studies of Elementary Chemisorption Reactions on an Activated Diamond (111) Terrace," J. Phys. Chem 98 (1994) 8128-8133.
20. Zhu, W., Pan, Z., Ho, Y., and Z. Man, "Impact-induced chemisorption of C_2H_2 on diamond (001) surfaces: a molecular dynamics simulation,"

Nuclear Instruments and Methods in Physics Research B **153** (1999) 213 – 217.

21. Hu, Y., and S. B. Sinnott, " A molecular dynamics study of thin - film formation via molecular cluster beam deposition: effect of incident species," Surface Science **526** (2003) 230-242.
22. Hagan, M. T., Demuth, H.B. and M.H. Beale, "Neural Network Design," (Martin Hagan, 2002).
23. Landau, L.J., J. G. Taylor (Eds), "Concepts for Neural Networks – A Survey," (London : Springer – Verlag c 1998).
24. Bakircioğlu, H., and T.Koçak, "Survey of random neural network applications", European journal of Operational Research **126** (2000) 319 – 330.
25. Lennox, B., Montague, G. A., Frith, A.M., Gent, C., and V.Bevan, " Industrial application of neural networks – an investigation," Journal of Process control **11** (2001) 497 – 507.
26. Konno, K., Kamei, D., Yokosuka, T., Takami, S., Kubo, M., and A.Miyamoto, "The development of computational chemistry approach to predict the viscosity of lubricants," Tribology International **36** (2003) 455-458
27. Pendharkar, P. C., " A threshold – varying artificial neural network approach for classification and its application to bankruptcy prediction problem," Computers and operations research **32** (2005) 2561-2582

28. Natale, C. D., Proietti, E., Diamanti, R. and A.D'Amico, « Modeling of APCVD-Doped silicon dioxide deposition process by a modular neural network," IEEE transactions on semiconductor manufacturing, **12** (1999) 109-115.
29. Han, S. and G. S. May, "Using Neural Network process models to perform PECVD Silicon Dioxide Recipe Synthesis via Genetic Algorithms", IEEE Transactions on Semiconductor Manufacturing, **10** (1997) 279-287.
30. Geisler, J. P., Lee, C. S. G. and G.S. May, "Neurofuzzy modeling of chemical vapor deposition process," IEEE Transactions on Semiconductor Manufacturing, **13** (2000) 210-215.
31. Hobday, S., Smith, R. and J. Belbruno, "Applications of neural networks to fitting interatomic potential functions," Modeling Simul. Mater. Sci.Eng. **7** (1999) 397 – 412.
32. Lorenz, S. Groβ, A. and M. Scheffler, "Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks," Chemical Physics Letters **395** (2004) 210-215.
33. Raff, L. M., Malshe, M., Hagan, M., Doughan, D.I., Rockley, M.G., and R.Komanduri, " *Ab initio* potential - energy surfaces for complex, multichannel systems using modified novelty sampling and feedforward neural networks," J.Chem.Phys. **122** (2005) 084104-1 - 084104-16.
34. Sumpter, B. G. and D. W. Noid, "Potential Energy surfaces for macromolecules. A neural network technique," Chemical Physics Letters **192** (1992) 455-462.

35. West, D., Dellana, S., and J.Qian. “Neural network ensemble strategies for financial applications,” Computers and Operations Research **32** (2005) 2543-2559.
36. Coyne, R. D., and A. G. Postmus, “Spatial applications of neural networks in computer-aided design,” Artificial intelligence in Engineering **5** (1990) 9-22.
37. Faller, W. E., and S. J. Schreck, “Neural Networks: Applications and ooportunities in Aeronautics,” Prog.Aerospace Sci. **32** (1996) 433-456.
38. Haile,. J. M., *Molecular dynamics simulation* elementary methods (New York: Wiley c 1992).
39. Rapaport, D. C., *The art of Molecular Dynamics simulation* (Cambridge University Press c 1995).
40. Koháry, K., Kugler, S., and I.László, “Molecular dynamics simulations of amorphous carbon structures,” Journal of Non-crystalline solids **227-230** (1998) 594-596.
41. Hu, Y., and S. B. Sinnott, “Molecular dynamics simulation of thin film nucleation through molecular cluster beam deposition: Effect of incident angle,” Nuclear Instruments and Methods in Physics Research B **195** (2002) 329 – 338.
42. Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. V., DiNola, A. and J.R.Hoak, “Molecular dynamics with coupling to an external bath,” J. Chem. Phys. **81** (1984) 3684-3690.

43. Komanduri, R. and L. M. Raff, “A Review on the molecular dynamics simulation of machining at the atomic scale,” Proc. Instn. Mech. Engrs. Part B 215 (2001) 1639-1672.
44. Retrieved May 19, 2005 from the world wide web: *http://vv.carleton.ca/~neil/neural/neuron-a.html*

VITA

Sivakumar Mathanagopalan

Candidate for the Degree of

Master of Science

Thesis: A NEURAL NETWORK AND MOLECULAR DYNAMICS APPROACH FOR EVENT PROBABILITY PREDICTION DURING CHEMICAL VAPOR DEPOSITION OF A CARBON ATOM ON DIAMOND (111) SURFACE

Major Field: Mechanical Engineering

Biographical:

Personal: Born in Chennai, Tamil Nadu, India on May 23, 1981, the son of Mr. B. Mathanagopalan and Mrs. S. Saradha

Education: Received Bachelor of Engineering degree in Mechanical Engineering from Madurai Kamaraj University, TamilNadu, India in May, 2002; completed requirements for the Master of Science degree at Oklahoma State University in July 2005.

Experience: 1. Graduate Research Assistant in School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, Oklahoma, August 2004 – May 2005.

2. Graduate Teaching Assistant in School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, Oklahoma, January 2003 – December 2004.