



# NEURAL NETWORKS IN CHEMICAL REACTION DYNAMICS

LIONEL M. RAFF,  
RANGA KOMANDURI,  
MARTIN HAGAN, AND  
SATISH T.S. BUKKAPATNAM

# Neural Networks in Chemical Reaction Dynamics

*This page intentionally left blank*

# Neural Networks in Chemical Reaction Dynamics

Lionel M. Raff, Ranga Komanduri,  
Martin Hagan, and Satish T.S. Bukkapatnam

OXFORD  
UNIVERSITY PRESS

OXFORD  
UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further  
Oxford University's objective of excellence  
in research, scholarship, and education.

Oxford New York  
Auckland Cape Town Dar es Salaam Hong Kong Karachi  
Kuala Lumpur Madrid Melbourne Mexico City Nairobi  
New Delhi Shanghai Taipei Toronto

With offices in  
Argentina Austria Brazil Chile Czech Republic France Greece  
Guatemala Hungary Italy Japan Poland Portugal Singapore  
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright © 2012 by Oxford University Press

Published by Oxford University Press, Inc.  
198 Madison Avenue, New York, New York 10016  
www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
electronic, mechanical, photocopying, recording, or otherwise,  
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

Neural networks in chemical reaction dynamics / Lionel M. Raff... [et al].  
p. cm.

Includes bibliographical references and index.

ISBN 978-0-19-976565-2 (hardcover : acid-free paper)

1. Chemical reactions—Data processing. 2. Neural networks (Computer science)  
I. Raff, Lionel M. II. Title.

QD501.N48 2011  
541'.390285632—dc22

2010054098

1 3 5 7 9 8 6 4 2  
Printed in the United States of America  
on acid-free paper

Dedicated with Admiration and Respect to the Memory of

Professor Ranga Komanduri

Oklahoma State University Regents Professor  
A. H. Nelson, Jr. Endowed Chair in Engineering  
Scholar, Teacher, Mentor, Scientist, and Engineer



*This page intentionally left blank*

# CONTENTS

Preface	xi
Acronyms	xiii

## 1. Fitting Potential-Energy Hypersurfaces

1.1. Introduction	1
1.2. Empirical and Semi-Empirical Potential Surfaces	2
1.3. <i>Ab Initio</i> Potential-Energy Surfaces (PESs)	3
1.4. Other Fitting Methods for Potential-Energy Surfaces	4
1.5. Neural Network (NN) Approach	7
1.6. Essential Steps in a Molecular Dynamics Simulations	7
1.7. Organization of the Monograph	8

## 2. Overview of Some Non-Neural Network Methods for Fitting *Ab Initio* Potential-Energy Databases

2.1. Introduction	13
2.2. Moving Shepard Interpolation (MSI) Methods	13
2.2.1. Required Input Data	13
2.2.2. MSI Method for Molecules with Four or Fewer Atoms	14
2.2.3. MSI Method for Molecules with More than Four Atoms	16
2.2.4. MSI Configuration Space Sampling	21
2.2.5. Applications and Results	23
2.3. Interpolative Moving Least-Squares Methods (IMLS)	24
2.3.1. General Method	24
2.3.2. Cutoff Function, Basis Sets, and Data Sampling	28
2.3.3. Applications and Results	30
2.4. Invariant Polynomial (IP) and Reproducing Kernel Hilbert Space (RKHS) Methods	31
2.4.1. Invariant Polynomial Methods	31
2.4.2. Applications and Results of IP Methods	33
2.4.3. Reproducing Kernel Hilbert Space (RKHS)	34
2.5. Hybrid Methods	37
2.5.1. Application to H <sub>3</sub> System	37
2.5.2. Application to the O( <sup>1</sup> D) + H <sub>2</sub> System	38
2.6. Neural Networks Applications to Reaction Dynamics	40

**3. Feedforward Neural Networks**

- 3.1. Introduction 42
- 3.2. Neuron Model 42
- 3.3. Network Architectures 44
- 3.4. Approximation Capabilities of Multilayer Networks 45
- 3.5. Training Multilayer Networks 48
- 3.6. Generalization (Interpolation and Extrapolation) 51
- 3.7. Data Preprocessing 53
- 3.8. Practical Aspects of NN Training Issues 54
  - 3.8.1. Database, Local Minima, Sampling Bias, Committees, and Derivatives 54
  - 3.8.2. Input Vector Optimization and Fitting Accuracy 55
- 3.9. Example Training Process (MATLAB) 60
- 3.10. The Combined Function Derivative Approximation (CFDA) NN Method 63
- 3.11. Combined Function Derivative Approximation Pruning 67
  - 3.11.1. Two-Layer Network Response 68
  - 3.11.2. Two-Layer Network Response 71
    - 3.11.2.1. Type-A Overfitting 71
    - 3.11.2.2. Type-B Overfitting 73
  - 3.11.3. Pruning Algorithm Summary 75

**4. Configuration Space Sampling Methods**

- 4.1. Introduction 77
- 4.2. Trajectory and Novelty Sampling (NS) Methods 78
- 4.3. Self-Starting Method Using Direct Dynamics (DD) 87
- 4.4. Configuration Sampling Using a Gradient Fitting Method 94

**5. Applications of Neural Network Fitting of Potential-Energy Surfaces**

- 5.1. Introduction 100
- 5.2. Near Equilibrium Structures—Vibrational State Studies 104
  - 5.2.1. The  $\text{H}_3^+$  Molecular Ion 104
  - 5.2.2.  $\text{H}_2\text{O}$ ,  $\text{HOOH}$ , and  $\text{H}_2\text{CO}$  105
- 5.3. CFDA Fitting—The  $\text{H} + \text{H}'\text{Br} \rightarrow \text{HBr} + \text{H}'$  and  $\text{H}_2 + \text{Br}$  Reactions 108
- 5.4. *cis-trans* Isomerization and N-O Dissociation Reactions of HONO 116
- 5.5. Gradient Sampling—Unimolecular Dissociation of  $\text{HOOH}$  to 2 OH 124
- 5.6. Unimolecular Dissociation of Vinyl Bromide ( $\text{H}_2\text{C} = \text{CHBr}$ ) 127
- 5.7. Non-Adiabatic Reactions:  $\text{SiO}_2 \rightarrow \text{SiO} + \text{O}$  and  $\text{SiO}_2 \rightarrow \text{Si} + \text{O}_2$  133
- 5.8. Generalized NN Representation of High-Dimensional Potential-Energy Surfaces 141

**6. Potential-Energy Surfaces Using Expansion Methods and Neural Networks**

- 6.1. Introduction and Overview of Expansion Methods 148

6.2. High-Dimensional Model Representation (HDMR) and NNs	151
6.3. Many-Body Expansions, Moiety Energy Approximations, and NNs	155
<b>7. Genetic Algorithm (GA) and Internal Energy Transfer Calculations Using Neural Network (NN) Methods</b>	
7.1. Genetic Algorithm (GA) Calculations Using NN Methods	165
7.1.1. Introduction	165
7.1.2. Brief Overview of GA	166
7.1.3. Application of NNs for GA Acceleration	167
7.1.4. Interatomic Potential Functions for Si	168
7.1.5. GA-NN Approach	170
7.1.6. Application to Fitting of PES for Si	173
7.1.7. Summary	181
7.2. Internal Energy Transfer Calculations Using NN Methods	181
7.2.1. Introduction to Internal Energy Transfer Calculations	181
7.2.2. Applications for Acceleration of IVR Calculations: H <sub>2</sub> O <sub>2</sub> System	183
7.2.2.1. NN for Conversion of Cartesian Velocities and Kinetic Mode Energies	183
<b>8. Empirical Potential-Energy Surfaces Fitting Using Feedforward Neural Networks</b>	
8.1. Fitting to <i>Ab Initio</i> Electronic Structure Data	190
8.1.1. Introduction	190
8.1.2. General Method	191
8.1.3. Application to Fitting an <i>Ab Initio</i> Database for Si <sub>5</sub> Clusters	193
8.2. Fitting Empirical Potentials to Vibrational Spectral Data	199
8.2.1. Introduction	199
8.2.2. Application to Macromolecules	200
<b>9. Neural Network Methods for Data Analysis and Statistical Error Reduction</b>	
9.1. Introduction	204
9.2. Interaction of Carbon (C <sub>2</sub> ) Dimer with Diamond—MD Simulations	205
9.3. Statistical Data Analysis and Results	205
9.4. Conclusions	212
<b>10. Other Applications of Neural Networks to Quantum Mechanical Problems</b>	
10.1. Solving the Molecular Vibrational Schrödinger Equation	215
10.1.1. Introduction and General Theory	215
10.1.2. Application to H <sub>2</sub> O	220
10.1.3. Summary	221

10.2. Prediction of High-Level Electronic Structure Energies from Hartree-Fock Energies	223
10.2.1. Introduction	223
10.2.2. Concepts and General Procedures	225
10.2.3. Illustrative Application to Vinyl Bromide Database	229
10.2.4. Application to Equilibrium Energies of Molecular Systems	233
10.2.5. Discussion and Evaluation of the Method	237
10.2.6. Summary	242

## 11. Summary, Conclusions, and Future Trends

11.1. Introduction	244
11.2. Other Methods for Obtaining PESs from <i>Ab Initio</i> Data	247
11.3. Configuration-Space Sampling Methods	247
11.4. Feedforward NN Fitting of <i>Ab Initio</i> PESs	248
11.5. Expansion Methods and NNs	249
11.6. Genetic Algorithm and IVR Calculations Using NN Methods	250
11.7. NN Methods for Parameter Determination of Empirical PESs	251
11.8. Combined Function Derivative Approximation (CFDA) NN Method	252
11.9. NN Methods for Data Analysis and Statistical Error Reduction	253
11.10. Other Applications of NNs to Quantum Mechanical Problems	254
11.11. Future Trends	255

References	257
About the Authors	275
Subject Index	279

## PREFACE

During the past decade, research efforts by a number of different groups have demonstrated that feedforward neural networks (NN) can play an important role in facilitating the development of potential-energy surfaces (PES) of sufficient accuracy to execute accurate molecular dynamics (MD), Monte Carlo (MC), quantum scattering calculations, determination of rovibrational energy levels within an accuracy of a few  $\text{cm}^{-1}$ , and perhaps photodissociation, if the excited-state electronic surfaces can be computed. Specifically, it has been found that NNs can greatly facilitate the use of potential energies and gradients obtained from first-principles electronic structure calculations to develop what are generally termed *ab initio* PESs for use in MD, MC, and quantum dynamics investigations. Likewise, NNs have been shown to be highly useful in making older methods, such as many-body expansions and genetic algorithms (GA), far more robust and powerful.

Neural network (NN) methods have been developed to greatly assist in the adjustment of parameters of empirical PESs to fit a database of some type. They have also been found highly useful in reducing the statistical error that is always present in MD studies. This has been accomplished by employing NNs to predict the results of trajectories without actually executing numerical integrations. When this is done, the results of a huge number of molecular trajectories can be obtained with far less computational effort. The result is a significant reduction in the statistical errors of the calculations and a corresponding increase of predictive accuracy. Most recently, it has been shown that NNs can be used to predict the results of high-level electronic structure calculations from lower-level Hartree-Fock calculations. This allows the computational bottleneck of executing the extensive *ab initio* electronic structural calculations to be substantially reduced.

While all these methods have been reported by several research groups in the leading journals of the field, there is no succinct compilation of all these results and methods that describes both the methods and the techniques involved in their application. This monograph provides clear descriptions of how NNs have been used to move MD investigations into the realm of *ab initio* calculations, which as recently as a few years ago were considered beyond the reach of the computational facilities available. The monograph also describes in detail how NNs can be effectively employed to execute all of the studies previously mentioned. In addition to the formal descriptions of the methods, example applications are provided to illustrate the power and limitations of NN methods along with a clear road map describing how to execute each of the studies.

The authors feel confident that this emerging technology will play an increasingly important role in MD, MC, and quantum mechanical studies of chemical reaction dynamics in the years to come.

The authors would like to thank the National Science Foundation (NSF) for continued support over the past 20 years of the work presented here. In particular, the authors thank the Manufacturing Processes and Machines Program of the Civil, Mechanical, and Manufacture Innovation (CMMI) Division of NSF [Drs. B. M. Kramer, J. Cao, J. Larsen-Basse, K. Srinivasan, K. Rajurkar, G. Hazelrigg]. The authors also thank the graduate students, in particular Milind Malshe, Rutuparna Narulkar, A. Pukrittayakamee, and H. Lu who made significant contributions as part of their Ph.D. dissertations. The atomistic simulation group at Oklahoma State University has had many graduate students who worked on various MD/MC projects over time in the past two decades. To all of them, and especially to Dr. Naga Chandrasekaran, the authors are grateful for their valuable contributions. The authors also thank Professor M. Rockley and Dr. P. M. Agrawal for their valuable contributions and Mrs. Sharon Green for her superb job of providing editorial assistance in the preparation of this monograph and several papers the authors have published in this area.

This monograph is based in large part on the recent work by the authors' interdisciplinary group on neural networks in chemical reaction dynamics. Consequently, no effort was made to include a comprehensive review of the literature on chemical reaction dynamics or neural networks. Instead, only references directly related to the subject matter are covered to limit the scope of the monograph. Also, references covering the entire manuscript are given toward the end of the monograph instead of at the end of each chapter to avoid duplication and save space.

Special thanks are extended to Professor Tucker Carrington Jr. and Dr. Sergei Manzhos for reviewing portions of the manuscript and for providing cogent suggestions for improving it. We also thank them for granting permission to use material from their published work. We also wish to thank Professors J. Behler and M. Parrinello and Drs. B. G. Sumpter and D. W. Noid for allowing us to use figures and tables from their papers as well as for providing helpful discussion of various points.

Finally, the authors sincerely thank various publishers for generously granting permission to use material the authors have published in their journals, including the *Journal of Chemical Physics*, *Physical Review*, *Journal of Physical Chemistry*, and *Philosophical Magazine*. All the figures and tables used in the monograph from these journals are duly acknowledged with a reference number giving details of the article in the reference section and a statement stating "Reprinted with the Permission from the publisher (name)" at the end of each figure caption and table caption.

## ACRONYMS

BFGS	Broyden-Fletcher-Goldfarb-Shannon Method
BP	Behler - Parrinello Method
CASPT2	Second-Order Complete Active Space Calculations
CFDA	Combined Function Derivative Approximation
CNP	Complete Nuclear Permutation
CPU	Central Processing Unit
CVD	Chemical Vapor Deposition
DAC	Density Adaptive Cutoff
DD	Direct Dynamics
DD/MD/NS/NN	Direct Dynamics/Molecular Dynamics/Novelty Sampling/Neural Network Method
DFT	Density Functional Theory
EAM	Embedded-Atom Method
EMS	Efficient Microcanonical Sampling
FRC	Fixed Radius Cutoff
GA	Genetic Algorithm
HDMR	High-Dimensional Model Representation
HF	Hartree-Fock
IMLS	Interpolating Moving Least Squares
IP	Invariant Polynomial
IR	Infrared
IRP	Intrinsic Reaction Path
IVR	Intramolecular Vibrational Relaxation
LBS	Large Basis Sets
L-IMLS	Local Interpolative Moving Least Squares
LM	Levenberg-Marquardt
LSTH	Liu, Siegbahn, Truhlar, Horowitz
MAE	Mean Absolute Error
MC Simulation	Monte Carlo Simulation
MD Simulation	Molecular Dynamics Simulation
MD/NS/NN	Molecular Dynamics/Novelty Sampling/Neural Network Method
ME	Moiety Energy
MEAM	Modified Embedded-Atom Method
MEP	Minimum-Energy Path
MLP	Multi-Layer Perception

## Acronyms

MNS	Modified Novelty Sampling
MSI	Moving Shepard Interpolation
NN	Neural Network
NS	Novelty Sampling
PES	Potential-Energy Surface
RBFFNN	Radial Basis Function Neural Network
RC	Redundant Coordinates
RHF	Restricted Hartree-Fock
RKHS	Reproducing Kernel Hilbert Space
RPE	Relative Percentage Error
SBS	Small Basis Sets
RMSE	Root Mean Square Error
SNN	Symmetric Neural Network
SVD	Singular Value Decomposition
UNN	Unsymmetric Neural Network
VB	Vinyl Bromide

# Neural Networks in Chemical Reaction Dynamics

*This page intentionally left blank*

# 1

## FITTING POTENTIAL ENERGY HYPERSURFACES

### 1.1. INTRODUCTION

Molecular dynamics (MD) and Monte Carlo (MC) simulations are the two most powerful methods for the investigation of dynamic behavior of atomic and molecular motions of complex systems. To date, such studies have been used to investigate chemical reaction mechanisms, energy transfer pathways, reaction rates, and product yields in a wide array of polyatomic systems. In addition, MD/MC methods have been successfully applied for the investigation of gas-surface reactions, diffusion on surfaces and in the bulk, membrane transport, and synthesis of diamond using chemical vapor deposition (CVD) techniques. The structure of vapor deposited rare gas matrices has been studied using trajectories procedures.

If the chemical reaction of interest contains three atoms or fewer, various types of quantum and semiclassical calculations can be brought to bear on the problem. These methods include wave packet studies, close-coupling calculations at various levels of accuracy, and S-matrix theory. Several excellent review articles have been published describing the principal techniques and problems involved in conducting MD studies; the reader may wish to consult these as background material for this discussion.<sup>1-7</sup>

With the advent of relatively inexpensive, powerful personal computers, MD/MC simulations have become routine. Once the potential-energy hypersurface for the system has been obtained, the computations are straightforward, though time-consuming. In the majority of cases, the computational time required is on the order of hours to a few days. However, the accuracy of these simulations depends critically on the accuracy of the potential hypersurface used.

The major problem associated with MD/MC investigations is the development of a potential-energy hypersurface whose topographical features are sufficiently close to those of the true, but unknown, surface that the results of the calculations are experimentally meaningful. Once the potential surface is chosen or computed, all the results from any quantum mechanical, semiclassical, or classical scattering or equilibrium calculation are determined. The only purpose of the MD calculations is

to ascertain what these results are. Therefore, the most critical part of any MD/MC study is the development of the potential-energy hypersurface and the associated force field. Surprisingly, this is often the portion of the investigation to which least effort is devoted. This situation arises because of the inherent difficulty associated with this part of the overall problem.

## 1.2. EMPIRICAL AND SEMI-EMPIRICAL POTENTIAL SURFACES

When four or more many-electron atoms are present, global potentials have usually been obtained using empirical methods that rely heavily upon ad hoc parametrized functional forms suggested by physical and chemical considerations. The parameters contained in these forms are generally fitted to equilibrium thermochemical, spectroscopic, and structural data. If experimental activation energies are known, the barrier heights predicted by the empirical potential are adjusted to these values. Such methods have been employed by numerous investigators since the mid-1970s.

Many engineering applications of MD/MC studies, including machining, indentation, uniaxial tension, and tribological simulations to date have relied on empirical potentials, such as pairwise sums of Morse<sup>8,9</sup> or Lennard-Jones potentials<sup>10</sup> for metals. Baskes and his colleagues have developed an embedded-atom method (EAM)<sup>11</sup> and a modified embedded-atom method (MEAM)<sup>12</sup> for application to fcc and bcc metals, and even to hcp metals. For covalently bonded, semiconductor materials, such as Si, Ge, and diamond many-body potentials, such as Tersoff's potential,<sup>13</sup> the Bolding-Anderson potential,<sup>14</sup> and the Brenner potential<sup>15</sup> have been developed. These potentials provide approximate descriptions of most materials at or near their equilibrium configuration. However, when the lattice atoms are in configurations far removed from equilibrium, the accuracy of any empirical potential can be expected to decrease significantly. Such loss of accuracy also occurs in many chemical reactions that sample regions of configuration space that are far removed from equilibrium.

The limitations and liabilities associated with the use of empirical potential surfaces in MD calculations may be summarized as follows:

1. Since the ad hoc empirical functional form is simply concocted with little or no theoretical foundation, it is highly unlikely that it will accurately predict the correct experimental force field.
2. The parameters contained in an empirical potential are generally adjusted to fit equilibrium and stationary-point data, such as bond lengths, bond angles, vibrational frequencies, reactant and product equilibrium energies, elastic moduli, sublimation energies, and potential barrier heights. As a result, the potential usually fits the stationary points on the surface acceptably well. However, MD and MC simulations are primarily concerned with the energies and forces at nonstationary points. It is very unlikely that these quantities are accurately represented in configurations far removed from equilibrium.

3. If the database contains data from *ab initio* quantum mechanical calculations at nonstationary points as well as experimental data, the problem of fitting the empirical parameters becomes very difficult. Such fitting must be accomplished using iterative adjustment methods that often require months of effort. For example, nine months of laborious iterative fitting was required to obtain an acceptable fit of the 100 parameters contained in the empirical function used to fit the global potential surface for vinyl bromide<sup>16 a,b</sup>.
4. Once an empirical potential is developed for a particular complex system, there is no straightforward means by which it can be improved. Any alteration of the potential requires that the entire fitting process be repeated.
5. Empirical potentials are specific for the system under consideration. The procedures used to obtain such potentials cannot be easily automated and the methods applied to all systems. Every system must be treated individually. Therefore, the research is tedious and labor intensive.

Each of the problems outlined here is essentially intractable. Because of this, the accuracy of MD simulations on complex systems has, for practical purposes, reached its limit. If we wish to advance beyond this natural limit for empirical and hybrid potentials, we must develop ways to effectively utilize quantum mechanical methods that have the power to produce more accurate surfaces and force fields.

### 1.3. AB INITIO POTENTIAL-ENERGY SURFACES (PESs)

In principle, *ab initio* methods can produce much more accurate potential-energy surfaces for molecular dynamics. The fundamental problem is that electronic structure calculations become computationally intractable when the number of atoms and electrons present in the system becomes large. Nevertheless, excellent results have been obtained for numerous three- and four-body systems. In this section, we present a brief review of some of the systems studied and the methods employed to obtain the required *ab initio* potentials.

In many investigations, an empirical or semi-empirical potential-energy surface (PES) is employed to represent the force field of the system under investigation. Such potentials generally yield only qualitative or semiquantitative descriptions of the system dynamics. Empirical potential surfaces can be significantly improved by fitting the chosen functional form for the potential to the force fields obtained from trajectories using *ab initio* Car-Parrinello<sup>17</sup> molecular dynamics simulations. This method has been employed by several research groups.<sup>18-21</sup>

When the system of interest contains four atoms, usually it is no longer possible to conduct electronic structure calculations in all regions of configuration space. Consequently, importance sampling procedures must be employed to identify the critical regions of configuration space. In addition, interpolation between the computed points on the surface becomes increasingly difficult.

If the system under investigation is not too large and complex, *ab initio* electronic structure calculations offer a route to obtaining extremely large databases, which can be employed to obtain the PES for the system. Malshe et al.<sup>22</sup> have recently developed a method that reduces or eliminates many of the problems associated with fitting empirical potentials particularly when some of the parameters are made functions of the system configuration. The method completely obviates the problem of selecting the form of the functional dependence of the parameters upon the system's coordinates. This form is, in effect, determined automatically by a neural network (NN). This use of NNs will be discussed in more detail in Chapter 8.

#### 1.4. OTHER FITTING METHODS FOR POTENTIAL-ENERGY SURFACES

Several research groups have employed methods that obviate the need to select arbitrary functional forms for the potential surface. These methods attempt to accurately sample the configuration space of the system and then fit the resulting database of *ab initio* energies using some generalized numerical procedure. Five such methods that have been frequently employed are (i) moving interpolation techniques,<sup>23-25</sup> (ii) reproducing kernel Hilbert space (RKHS),<sup>26,27</sup> (iii) interpolating moving least squares (IMLS),<sup>28-33</sup> (iv) expansions in terms of invariant polynomials (IP),<sup>34-39</sup> and (v) neural network (NN) methods.<sup>40-50</sup> The following discussion briefly describes these methods. The first four are discussed in more detail in Chapter 2 while the remainder of this monograph is devoted to an extensive discussion of various applications of NN methods.

Ischtwan and Collins<sup>23</sup> have developed a moving interpolation technique in which the potential energy in the neighborhood of any point is approximated by a Taylor series expansion using inverse bond length coordinates as the expansion variables. The overall method is usually denoted as a moving Shepard interpolation (MSI). Initially, a set of system configurations along the reaction path is selected by using chemical intuition. *Ab initio* electronic structure calculations of the system energy and gradients in these configurations are then executed. The data thus produced are employed to obtain the set of Taylor series expansions. A Shepard method<sup>24</sup> is used to effect the required interpolations. This procedure expresses the potential at any configuration as a weighted average of the Taylor series about all  $N$  points in the data set, or alternatively, those points within a specified cutoff radius. Subsequently, the results are iteratively improved by computing trajectories on the Taylor series fitted surface and recording the internal coordinates at a series of successive configurations encountered in the trajectories. In part, these new points are added to the data set according to a weight factor that is determined by the relative density of points in the data set in the region of the new points. For systems that are undergoing two-center, bond dissociation or formation reactions, the fitting error with a test set of *ab initio* energies is usually found to be in the range of from 0.0010 eV to 0.031 eV depending upon the complexity of the system. For an overview of this method and results, the review by Collins may be consulted.<sup>25</sup>

The second method for obtaining *ab initio* potential surfaces for three- and four-body systems employs the criterion of reproducing kernel Hilbert space (RKHS) as the means to effect the required fitting of the *ab initio* energies. Ho et al.<sup>26</sup> have employed this procedure to investigate the  $\text{O}(^1\text{D}) + \text{H}_2 \rightarrow \text{OH} + \text{H}$  reaction. When 1280 *ab initio* data points were employed to define the surface, the fitting yielded a surface whose absolute root mean square error compared to *ab initio* energies was 0.0131 eV. Pederson et al.<sup>27</sup> have used the same method to study the dynamics of the  $\text{N}(^2\text{D}) + \text{H}_2$  reaction.

Maisuradze et al.<sup>28</sup> have introduced the third method, namely, an interpolating moving least-square (IMLS) method to effect the fitting between the computed *ab initio* points. When the method is unrestricted, the least-square coefficients are obtained from the solution of a large matrix equation that must be solved repeatedly during a trajectory study. In this form, the required computational time increases with  $NM^2$ , where  $N$  is the number of data points to be fitted and  $M$  is the number of basis functions used in the linear combination that provides the fit. Guo et al.<sup>29</sup> have evaluated the IMLS method using the analytic surface of Kuhn et al.<sup>51</sup> for simple O-O bond rupture of HOOH. With  $N = 300$ , the results showed the statistical fitting errors to lie in the range of 0.085 eV to 0.171 eV. Kawano et al.<sup>30</sup> have compared the accuracy of the IMLS method with that of the MSI procedure for the HOOH bond scission reaction. With  $N = 6489$ , the root mean square errors for IMLS and MSI were 0.0468 eV and 0.0202 eV, respectively. Guo et al.<sup>31</sup> have recently reported IMLS results for the unimolecular dissociation reaction of  $\text{H}_2\text{CN} \rightarrow \text{H} + \text{HCN}$ . Their results show root mean square fitting errors that vary between 0.0055 eV and 0.0290 eV.

The major problem with the IMLS method is the extremely large computational time required for its execution. When more than four atoms are involved, the computation time increases rapidly due to the  $M^2$  factor. The time also increases linearly with the number of *ab initio* data points employed. Guo et al.<sup>31</sup> have reported a computational time of about 100 minutes for one  $\text{H}_2\text{CN}$  dissociation trajectory lasting 2.5 ps when 830 data points are employed.

The computational difficulties associated with IMLS methods have recently been addressed by Dawes et al.<sup>32a</sup> and by Guo et al.<sup>32b</sup> By converting the IMLS method to what is essentially a modified Shepard method,<sup>25</sup> the large computational bottleneck present in IMLS calculations is avoided. The authors<sup>32</sup> term this hybrid method a “local IMLS” or L-IMLS method. To date, this hybrid procedure has been applied only to molecules containing four or fewer atoms undergoing a single, two-center, bond dissociation reaction. Ishida and Schatz<sup>33a</sup> have also proposed a procedure that combines the IMLS and Shepard methods. This combined method has been applied to the investigation of the three-body reaction dynamics of the  $\text{O}(^1\text{D}) + \text{H}_2$  reaction.<sup>33b</sup>

Bowman, Braams, and co-workers<sup>34–39</sup> have developed a different type of least-squares fitting procedure that utilizes least-squares fitting of expansions in terms of polynomials of functions of the interatomic distances to fit *ab initio* databases obtained from electronic structure calculations. The expansion polynomials are

totally symmetrized so that they are invariant to the exchange of any two identical atoms. For this reason, the procedure is generally called an “invariant polynomial” (IP) method. The IP method has been applied to an impressive set of relatively complex reactions of molecules that contain between five and seven atoms. In most cases, the database to which the IP expression is fitted contains more than 20,000 points.

The fifth method for obtaining accurate analytic fits to databases obtained from *ab initio* electronic structure calculations involves the use of NNs. Neural networks provide a powerful and robust method for surface fitting and other related tasks associated with reaction dynamics investigations. The remainder of this monograph is devoted to an exposition of these NN methods.

In order to avoid the extremely difficult task of interpolating *ab initio* electronic structure data, some investigators have employed a method known as “direct dynamics” (DD). In this approach, trajectories are computed by direct calculation of the force field at each integration point using some *ab initio* quantum mechanical method. Because of the huge number of computations required for each trajectory, the *ab initio* method chosen must usually be some form of density functional theory (DFT). For example, it may be necessary to follow the dynamics for several picoseconds (ps). With an integration step size on the order of a femtosecond (fs), 1,000 or more integration steps will be required. If the forces are computed at four points in each step, something on the order of 4,000 gradients will have to be computed for each trajectory. As a result, only a very limited number of trajectories can be obtained.

As the system under investigation increases in size, the computational requirements of DD will quickly overwhelm the available computational resources. The basic problem is that all the information obtained about the potential-energy surface and the corresponding force field during the integration of the trajectory is discarded after the completion of the calculation. Therefore, each subsequent trajectory fails to profit from all the computational effort expended in obtaining the previous trajectories. The conventional wisdom has been that *ab initio* MD calculations for complex systems containing five or more atoms with several open reaction channels are presently beyond our computational capabilities. The rationales for this view are (a) the inherent difficulty of high-level *ab initio* quantum calculations on complex systems that may make numerous, large-scale computations impossible, (b) the large dimensionality of the configuration space for such systems that makes it necessary to examine prohibitively large numbers of nuclear conformations, and (c) the extreme difficulty associated with accurate interpolation of numerical data obtained from electronic structure calculations when the dimensionality of the system is nine or greater. In subsequent chapters, we present general NN methods to handle each of these difficulties for many complex systems. Furthermore, as our computational resources increase, the range of systems that can be so treated will expand without the need to significantly modify the methods. Therefore, the theoretical approach is robust and powerful.

## 1.5. NEURAL NETWORK (NN) APPROACH

In this monograph, we present an integrated approach for obtaining *ab initio* quantum mechanical potential-energy surfaces (PES) and force fields for use in MD simulations for systems of five or more atoms in which several reaction channels may be energetically open. The method involves (a) electronic structure calculations of energies and force fields for an ensemble of atoms whose interatomic distances are such that all atoms fall within a previously specified distance from some central point, generally called the cutoff radius; (b) the implementation of sampling methods that utilize MD calculations and novelty sampling to permit the regions of configuration space that are important in the dynamics to be accurately identified; and (c) the use of NNs, early stopping, and regularization methods to provide rapid and accurate fitting of the *ab initio* database and convenient tests for convergence that do not require MD calculation of dynamical results.

## 1.6. ESSENTIAL STEPS IN MOLECULAR DYNAMICS SIMULATIONS

For pedagogical purposes, we may divide an MD investigation into several distinct parts. The division used here is not unique, but it serves the purpose of guiding the discussion that follows. The essential steps of any MD study may be considered to be the following:

*Step 1:* A convenient coordinate system must be chosen that leads to a set of coupled, ordinary differential equations whose solution determines the nature of a molecular trajectory for a given set of initial conditions. This step concludes with the preparation of computer codes that perform numerical integrations to solve the coupled set of differential equations.

*Step 2:* The molecular motions during a trajectory are determined by the force field in which the atomic particles move. This force field is obtained from the gradients of the potential-energy surface at each integration point. The development of the PES is, therefore, a crucial step in the MD study.

*Step 3:* Appropriate methods must be devised to determine when to terminate trajectories and how to treat the final results to obtain the experimental quantities of interest in the investigation. Since the calculations are classical rather than quantum mechanical, if the quantities of interest involve quantization of the final states, methods to convert the final classical results to corresponding quantum results with reasonable accuracy must be developed.

*Step 4:* Since experimental observations are usually ensemble averages of the results of an extremely large number of molecular collisions that is often on the order of the Avogadro number ( $10^{23}$ ), an MD investigation must incorporate appropriate statistical averages over the same variables averaged in the corresponding experiments. This is generally accomplished by the computation of a number of molecular trajectories in which the initial conditions are chosen randomly from the

appropriate distribution functions for those variables over which the investigator seeks to average. Obviously, it is computationally impossible to compute  $10^{23}$  trajectories, so these averages must of necessity be over much smaller sets that usually comprise from  $10^2$  to  $10^4$  trajectories. This naturally leads to the problem of estimating the statistical error present in the calculations and the development of methods that reduce this error as much as possible.

Although each of these steps is important in an MD investigation, the most important and the most difficult step in the execution of accurate MD, MC, semi-classical, or quantum scattering calculation is Step 2, the development of a sufficiently accurate PES. This is the case because once the PES is in hand, all of the dynamics for a given set of experimental conditions are determined. The MD, MC, or other scattering calculations are simply computational procedures used to discover the nature of the scattering or reaction dynamics that have already been determined by the formulation of the PES.

During the past decade, research efforts by a number of different groups have demonstrated that feed-forward, neural networks (NN) can play an important role in facilitating the development of PESs of sufficient accuracy to execute accurate MD, MC, or quantum calculations. Specifically, it has been found that NNs can greatly facilitate the use of potential energies and gradients obtained from first-principles electronic structure calculations to develop what are generally termed *ab initio* PESs for use in MD, MC, and quantum dynamics investigations. Likewise, NNs have been shown to be highly useful in making alternative methods to derive empirical PES, such as many-body expansions and genetic algorithms (GA), far more robust and powerful. Methods have been developed to greatly assist in the adjustment of parameters of empirical PESs to fit a database of some type.

In addition to facilitating the development of PESs, NNs have been found to be highly useful in Step 4 where one seeks to reduce the statistical error that is always present in MD studies. This has been accomplished by employing NNs to predict the results of trajectories without actually executing the numerical integrations of Step 1. When this is done, the results of a huge number of molecular trajectories can be obtained with far less computational effort. The result is a significant reduction in the statistical errors of the calculations and a corresponding increase of predictive accuracy. This procedure permits far more detailed investigation of the dependence of various experimentally measured properties on the independent variables of the process under examination.

## 1.7. ORGANIZATION OF THE MONOGRAPH

The objective of this monograph is to provide the reader with a brief but clear presentation of NNs along with descriptions of how they can be used to move MD investigations into the realm of *ab initio* calculations, which as recently as a few years ago were considered to be beyond the reach of present computational facilities. The discussion also describes in detail how NNs can be effectively employed to

study non-adiabatic chemical reactions on *ab initio* potentials, obtain PESs for large ensembles of atoms that are invariant to permutations of identical atoms, enhance GA and energy transfer calculations, render many-body expansions far more robust and powerful, facilitate quantum scattering calculations, predict the results of trajectories without actually integrating the equations of motions, obtain parameters of empirical PESs from *ab initio* data or vibrational spectra, and finally how NNs can play a key role in the determination of electronic structure energies and gradients from high-level computations using primarily calculations conducted at the Hartree-Fock (HF) level of theory.

The monograph is organized into 11 chapters. Chapter 1 begins with a brief introduction to various methods of fitting potential energy surfaces. They include multilayer perceptron neural networks; empirical and semi-empirical potential surfaces; *ab initio* potential-energy surfaces; other fitting methods, namely, (a) moving interpolation techniques, (b) reproducing kernel Hilbert space (RKHS), and (c) interpolating moving least squares (IMLS); (d) invariant polynomials (IP), and the neural network (NN) approach. The chapter ends with four essential steps in MD simulations and the organization of the monograph.

Chapter 2 provides a reasonably complete description of MSI, IMLS, and IP methods along with the presentation of typical results of applications of these methods. In addition, some hybrid methods that combine these techniques are also presented and typical results given. Results from the application of the RKHS method are presented.

Chapter 3 provides an introduction to the standard NN methods. It concentrates on the most common neural network architecture, namely, the multilayer perceptron (MLP). It describes the basics of this architecture, discusses its capabilities, and shows how it has been used on several different chemical reaction dynamics problems. This is followed by a more specific discussion of how NN methods can be employed in chemical reaction dynamics. For readers already familiar with NNs, parts of this chapter will provide a quick and easy-to-read starting material for their application to important problems in chemical reaction dynamics. Readers unfamiliar with NNs will find Chapter 3 highly useful as a means to bring themselves quickly to the operational level of the use of NNs. Finally, Chapter 3 takes the reader to the edge of the research frontiers in this area by presentation of a powerful new type of NN training algorithm that fits not only a function but also its gradient. This type of NN training is called the “Combined Function Derivative Approximation” (CFDA). Since it is the gradient of the potential energy that is of importance in molecular dynamics (MD) simulations, the CFDA promises to be a major player in future applications of NNs to chemical reaction dynamics. The use of the CFDA leads to a new type of overfitting that is not seen when NNs are employed in their usual form. Chapter 3 concludes with a discussion of this type of overfitting and “pruning” methods to permit its elimination.

In any fitting method, the procedures used to sample the configuration space of the system are of critical importance. In Chapter 4, we describe the theory and operational details of the three most powerful methods associated with configuration space

sampling, namely, (i) trajectory and novelty sampling (NS) methods, (ii) initiation using direct dynamics (DD), and (iii) configuration sampling using a gradient fitting method. These descriptions are made clear and instructive by numerous illustrations from the literature of actual applications to real chemical systems.

Chapter 5 is devoted to an extensive discussion of actual examples of NN applications to chemical reaction dynamics. The chapter begins with an overview of the applications reported to date. This overview is then followed by a series of illustrative examples of increasing complexity beginning with an NN method for the *ab initio* computation of the molecular vibrational levels of the  $\text{H}_3^+$  molecular ion. The dual NN methods for obtaining highly accurate *ab initio* molecular vibrational levels, recently developed by Manzhos et al.<sup>47</sup> are then described and illustrated by application to  $\text{H}_2\text{O}$ ,  $\text{H}_2\text{O}_2$ , and  $\text{H}_2\text{CO}$ .

Chapter 5 then moves on to more complex systems involving chemical reactions. An application of the CFDA method to the gas-phase abstraction of exchange reactions of hydrogen reacting with HBr is described. It is shown that the CFDA method can achieve fitting accuracies on the order of  $1\text{ cm}^{-1}$  error over the entire configuration space that is important in the reaction dynamics. Applications to the *cis-trans* isomerization and N–O bond dissociation reactions of HONO are then described. An application of configuration space sampling using gradient fitting is provided by a discussion of the dissociation dynamics of hydrogen peroxide.

Chapter 5 culminates with detailed descriptions of some of the most complex systems yet investigated using purely *ab initio* methods, namely, the unimolecular dissociation of vinyl bromide,  $\text{H}_2\text{C}=\text{CHBr}$ , into six, simultaneously open reaction channels, and the non-adiabatic dissociation of  $\text{SiO}_2$  to either  $\text{Si} + \text{O}_2$  or to  $\text{SiO} + \text{O}$ , both of which are non-adiabatic processes involving three different potential surfaces, and the treatment of a large ensemble (64) of silicon atoms using an NN method that automatically incorporates complete permutational symmetry of identical atoms.

Expansion methods to obtain PESs were proposed in the 1980s (see, for example, Murrell et al.).<sup>52,53</sup> However, owing to the bottleneck of the arbitrary and unknown nature of the expansion functions, only very limited use of these methods has been made to date. With the advent of NNs, these expansion methods have now been empowered and promise to be powerful techniques for the future.

Chapter 6 describes in detail two methods, namely, a high-dimensional model representation (HDMR) and a many-body expansion with a moiety-energy (ME) approximation that have been developed to date that combine NNs with expansion methods to produce powerful techniques for PES surface fitting to *ab initio* databases. The first of these is a combination of the high-dimensional model representation (HDMR) method, recently discussed in a series of publications by Rabitz and co-workers,<sup>54–70</sup> and NNs that has been proposed and implemented by Manzhos and Carrington.<sup>48</sup> The second method employs NNs with a many-body expansion that has been developed by Malshe et al.<sup>71</sup> Both methods provide solutions to the combinatorial problem that plagues expansion methods and both methods produce excellent results in terms of fitting accuracy.

Genetic algorithms (GA) have been a standard method for obtaining fits to databases. However, when the database is large, they are plagued with a heavy computational burden due to the computational time required to repeatedly compute the objective function. By utilizing NNs to predict the value of the objective function without actual computation, the entire GA method has been rendered far more powerful. The first portion of Chapter 7 describes this combined GA-NN method with an example of its utilization for obtaining a PES for  $\text{Si}_5$  clusters.<sup>72</sup> The second portion of the chapter is devoted to a discussion how the use of NNs can greatly accelerate studies of intramolecular vibrational energy relaxation (IVR).<sup>73</sup>

Often it is useful to employ an analytical functional form for the PES. When this is done, it is always necessary to extensively parameterize the analytic function to allow accurate fitting to the database. The parameter fitting for arbitrary functional forms can be extremely laborious. This is particularly the case whenever it is desired to make one or more of the parameters functions of the instantaneous configuration of the system. Malshe et al.<sup>22</sup> have recently utilized NNs to greatly facilitate parameter fitting and to make it possible to obtain the near optimum functional dependence of the parameter values on the molecular configuration of the system. Chapter 8 describes these methods in detail. This chapter concludes with an application of NNs that permits the determination of empirical parameters from measured vibrational spectra for macromolecular systems.<sup>74</sup>

Since the final result of an MD trajectory is uniquely determined by the starting conditions of the trajectory—namely, orientation variables, impact parameter, relative translational energy, etc.—if a reasonable database exists that maps these initial starting conditions onto the final results of the trajectory, an NN can, in principle, be trained to predict the final results using an input vector containing the starting conditions for the trajectory. In effect, the NN replaces all the laborious numerical integrations required to compute the trajectory. If this can be done, then it becomes possible to compute  $N$  trajectories for a system, fit an NN to the results of these  $N$  trajectories, and use the NN to predict the results for a huge number of trajectories without having to perform any numerical integrations. These results can then be utilized to greatly enhance the statistical accuracy of the MD simulations with near negligible computational effort. Chapter 9 describes the details of a method that brings this concept to fruition (see Agrawal et al.).<sup>75</sup> An example application of the method to the reactions of a  $\text{C}_2$  dimer with an activated diamond (100) surface is described as an illustration of the method.

Chapter 10 presents several applications of NNs to problems arising in quantum mechanics. The first of these describes the manner in which radial basis function neural networks (RBFNN) can be employed to solve the molecular vibrational Schrödinger equation to obtain wavefunctions and associated vibrational energies for multiple levels in a single calculation. In effect, an RBFNN is employed as an approximate wave function in a formalism in which the neurons of the RBFNN play the role of basis functions in the usual expansion methods for solution of the problem. Details of the method developed by Manzhos and Carrington<sup>76,77</sup> are described in the first section of Chapter 10.

The second section of Chapter 10 describes an application of NNs to simplify the electronic structure calculations required to obtain *ab initio* databases. As such, the method has the potential to become one of the most important applications of NNs to chemical reaction dynamics.

The biggest bottleneck to the execution of *ab initio* computation of reaction dynamics for complex systems is the extremely large computation times required to execute electronic structure calculations at a sufficiently high level of accuracy to obtain the extensive databases required to characterize complex chemical processes. This is, without doubt, the limiting factor in the execution of such investigations.

Malshe et al.<sup>78</sup> have recently found that because the Hartree-Fock (HF) energies are highly correlated with the energies obtained from higher-level electronic structure calculations, it is possible to accurately predict the results of higher-level computation from HF results. This finding implies that large databases can be computed at high level by computing only a small subset of the required database at high level, training an NN with the results, and then using the trained NN to predict the higher-level results from the HF energies. If this method proves to be robust, this will enormously expand the complexity of systems than can be investigated using purely *ab initio* methods. Also discussed in this section are analogous studies by Balabin and Lomakina<sup>79</sup> that demonstrate that DFT results with large basis sets can be obtained from DFT calculations that employ much smaller basis set by using a trained NN.

Chapter 10 concludes with a discussion of results obtained by several research groups that demonstrate how the accuracy of electronic structure calculations of equilibrium properties can be significantly improved by employing NNs trained by fitting computational results to experimental data.

Finally, *Neural Networks in Chemical Reaction Dynamics* concludes in Chapter 11 with a summary of the various applications of NNs related to chemical reaction dynamics that are covered in this monograph posing challenges, identifying some opportunities and limitations, and outlining future trends. The authors strongly believe that the approaches presented in this monograph using neural networks to develop potential-energy surfaces will play an increasingly important role in MD, MC, and quantum mechanical studies of chemical reaction dynamics in the years to come.

# 2

## OVERVIEW OF SOME NON-NEURAL NETWORK METHODS FOR FITTING *AB INITIO* POTENTIAL-ENERGY DATABASES

### 2.1. INTRODUCTION

In this chapter, we describe results obtained by five methods that have been employed to fit *ab initio* potential-energy. These methods are (i) moving or modified Shepard interpolation (MSI),<sup>23-25</sup> (ii) interpolative moving least squares (IMLS),<sup>28-33</sup> (iii) invariant polynomials (IP),<sup>36,37</sup> (iv) reproducing kernel Hilbert space (RKHS),<sup>26,27</sup> and (v) a hybrid method that combines MSI and IMLS methods.<sup>33a,33b</sup> The MSI and IMLS methods are described in some detail in the following. The IP and RKHS procedures are significantly more complex, and the reader is referred to the original papers for a more complete discussion of the details by which these methods are executed.

### 2.2. MOVING SHEPARD INTERPOLATION (MSI) METHODS

The moving or modified Shepard interpolation (MSI) method was developed primarily by Collins and co-workers.<sup>23-25</sup> The method employs electronic structure calculations to obtain the molecular potential energy at configuration points generated by an automated procedure. These data are then employed in a Shepard interpolation procedure<sup>24</sup> to obtain the potential energies of the system at points other than those in the database. This procedure involves expressing the local potential about each configuration point in a Taylor series expansion. The term “moving” in the title derives from the fact that the set of internal coordinates employed in the interpolation varies from point-to-point in the database.

#### 2.2.1. Required Input Data

Like all fitting methods, the MSI procedure requires the potential energy at a set of configuration points in the (3*N*-6) dimensional internal space of the system under

investigation. These energies are generally obtained using *ab initio* electronic structure methods at some level of accuracy.

In addition to the potential energies at each configuration point, the method also requires at least the first and second derivatives of the potential with respect to the coordinates being employed at each configuration point. These derivatives are needed to allow the local potential about a given configuration point in the database to be expressed in terms of a Taylor series expansion about that point. In principle, the MSI method may be extended to include third or fourth derivatives,<sup>80</sup> but in most applications, the expansions are truncated after the quadratic terms. In some cases, depending upon the electronic structure method being employed, these derivatives can be computed analytically. In other cases, they must be obtained numerically using finite differences in the neighborhood of each configuration point.

Since the MSI procedure requires at least first and second derivatives, the number of data items needed at each configuration point is significantly greater than just the energy. If there are  $N$  atoms in the system, there will be  $3N-6$  internal coordinates at each data point. Consequently, there will be  $3N-6$  first derivatives of the energy and  $(3N-6)(3N-5)/2$  second derivatives required since there are  $(3N-6)(3N-7)/2$  off-diagonal, mixed second derivatives and  $3N-6$  diagonal second derivatives. The total number of derivatives needed is, therefore,  $(3N-6)(3N-3)/2$ . If  $N = 4$ , 27 derivatives will be needed in addition to the energy, so that 28 input data will be required at each configuration point. For a five-atom system, this total is 55.

At first glance, this would seem to impose a severe computational bottleneck on the MSI methods, but this is often not the case. If the electronic structure method being employed permits first and second derivatives to be evaluated analytically, the computation of these quantities increases the required computational effort only slightly. On the other hand, if the electronic structure method does not allow analytic computation of the second derivatives, then there will be a significant increase in the required computational time. However, even in this case, there is a compensating feature of the MSI method. The availability of so much information about the topology of the PES at each data point usually means that the database required by the method to attain convergence will be much smaller than would be the case if only the energy at each point were available.

### 2.2.2. MSI Method for Molecules with Four or Fewer Atoms

When  $N \leq 4$ , the number of internal coordinates and the number of interparticle distances are the same. Consequently, the PES may be constructed using either the vector of interparticle distances,  $\mathbf{R} = \{R_1, R_2, \dots, R_{3N-6}\}$  or the vector of reciprocal interparticle distances,  $\mathbf{Z}$ , where  $Z_k = R_k^{-1}$ . Since potential energies generally vary with some power of the reciprocal distances, a more physically reasonable asymptotic behavior is obtained if one employs the  $Z_i$  as expansion variables rather than the  $R_i$ . This is the procedure used in all MSI applications.

At this point, it is assumed that the potential-energy database to which the analytic surface is to be fitted is available. The iterative procedure used to obtain this