The Journal of Chemical Physics

# Permutation invariant polynomial neural network approach to fitting potential energy surfaces

Cite as: J. Chem. Phys. **139**, 054112 (2013); https://doi.org/10.1063/1.4817187
Submitted: 19 June 2013 . Accepted: 17 July 2013 . Published Online: 06 August 2013

Bin Jiang, and Hua Guo

# Permutation invariant polynomial neural network approach to fitting potential energy surfaces

Bin Jiang and Hua Guo[a)]
*Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA*

(Received 19 June 2013; accepted 17 July 2013; published online 6 August 2013)

A simple, general, and rigorous scheme for adapting permutation symmetry in molecular systems is proposed and tested for fitting global potential energy surfaces using neural networks (NNs). The symmetry adaptation is realized by using low-order permutation invariant polynomials (PIPs) as inputs for the NNs. This so-called PIP-NN approach is applied to the H + $H_2$ and Cl + $H_2$ systems and the analytical potential energy surfaces for these two systems were accurately reproduced by PIP-NN. The accuracy of the NN potential energy surfaces was confirmed by quantum scattering calculations. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4817187]

## I. INTRODUCTION

As a natural consequence of the Born-Oppenheimer approximation, potential energy surfaces (PESs) play a central role in understanding molecular spectroscopy and reaction dynamics.[1] However, the development of accurate global PESs has in the past been hampered by difficulties associated with high-level *ab initio* calculations and with fitting the multi-dimensional potential energy function in a large configuration space. The commonly used approaches include direct spline,[2,3] many-body expansion (MBE),[1,4] reproducing kernel Hilbert space (RKHS),[5] modified Shepard interpolation (MSI),[6,7] permutation invariant polynomial (PIP),[8,9] interpolating moving least squares (IMLS),[10,11] and neural network (NN) methods.[12–14] Each method has its own advantages and disadvantages. In this publication, we will focus on the NN approach, which has shown great promise for fitting high-dimensional *ab initio* PESs for polyatomic systems in the gas phase and on surfaces.[15–33]

## II. SYMMETRY ADAPTATION IN NEURAL NETWORKS

Inspired by biological systems, artificial NNs have been used in many areas of science and technology,[34] including chemistry.[35] For our purpose here, NNs are essentially a (nonlinear) fitting tool that allows an accurate interpolation (but not extrapolation) of a pointwise representation of a global PES.[14] It has been established that NNs are capable of approximating an unknown real-valued multi-dimensional function up to arbitrary accuracy.[13] As illustrated in Fig. 1, an NN converts signals, namely a set of or input parameters ($G_i$), to a scalar output ($E$, i.e., potential energy) *via* one or more hidden layers of interconnecting neurons, which provide the flexibility in the fitting. The neurons are basically a set of nonlinear functions whose parameters are determined by "training," namely fitting of a large number of potential energy points in the relevant coordinate space. However, it is important to realize that the input does not have to be coordinates – they can be functions of coordinates.

Permutation symmetry is prevalent in molecular systems, and it sometimes is essential to accurately describe nuclear dynamics, particularly in floppy systems.[36] As a result, a key aspect in developing PESs is how to rigorously take the symmetry into account.[1,8,13] Since NNs recognize no symmetry, additional measures have to be taken in order to adapt the symmetry. Internal coordinates, such as bond lengths and bond angles, are commonly used as the input set of NNs,[14,29] but they often do not contain the full permutation symmetry of the system. A simple solution for small systems is to train NNs to learn the symmetry by duplicating the data set according to the symmetry.[24,29,30] But the size of the training data set could increase significantly and the symmetry is not strictly preserved.[13] A better way is to order the coordinates before fitting and use symmetry after the fitting to recover symmetrically equivalent points not included in the fit.[32] However, it is hard to avoid discontinuity at high symmetry points, which may cause problems for classical trajectory calculations.

In principle, permutation symmetry can be explicitly included in the NN fitting of PESs in two ways. One can, for example, construct symmetric neurons in the first hidden layer of the NN that preserves the full symmetry of the system. The corresponding potential energy will be automatically symmetrized. This scheme has been used in the work of Prudente *et al*. for the $H_3^+$ PES.[18] In a related work recently proposed by Nguyen and Le,[31] modifications were made on the weights of the first hidden layer leading to a smaller number of NN parameters. The disadvantages of these implementations are apparent: a significant effort is required to design a symmetric NN structure for a particular molecule of interest and the generalization is difficult.[31] Alternatively, one can prepare symmetrized coordinates as the input set as done by Gassner *et al*.[17] Obviously, these coordinates have to be carefully chosen because they are system dependent. It is worthwhile to note that several symmetrized functions have been proposed based on Fourier terms for molecule surface interactions, where symmetry exists for the translational motion

a)Author to whom correspondence should be addressed. Electronic mail: hguo@unm.edu

 

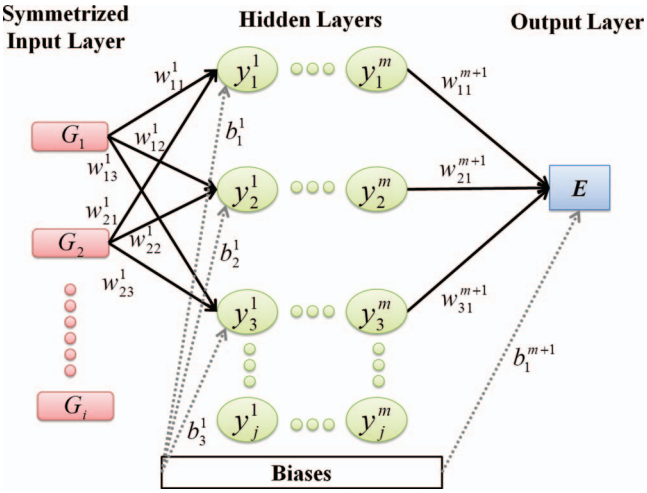


FIG. 1. Structure diagram of feed-forward neural networks with finite hidden layers.

parallel to the surface and rotational motion along the surface normal.[23,25] In addition, Behler and Parrinello have suggested to represent the total energy of the system as a sum of each atomic energy, which is respectively given by an individual neural network.[26] In their approach, the symmetry functions for each atom are constructed as many-body functions depending on all atomic positions inside a cutoff sphere. Since this method is designed for high-dimensional PES, its evaluation is more demanding because a specific NN structure is needed for each atom.[13]

We present here a new, simple, rigorous, and systematic scheme for defining symmetry functions for NN fitting of PESs, which is particularly suitable for gas-phase molecules. To this end, the idea of permutation invariant polynomials (PIPs)[8,9] is adapted. In the first step, we define a set of monomials, in terms of single-valued functions of internuclear distances, which provide a more straightforward representation for adapting permutation symmetry than internal or other coordinate systems. For example, the Morse like variables used in the PIP fitting approach of Bowman and co-workers[8,9] can be used,

$$p_{ij} = \exp(-a r_{ij}), \tag{1}$$

where $a$ is an adjustable parameter, $r_{ij}$ and $p_{ij}$ represent the bond length between the $i$th and $j$th atoms and the corresponding monomial. The use of the internuclear distances, rather than internal coordinates, is due to the ease in implementing permutation symmetry. Other possible choices of the monomials include inverse and logarithmic functions, as illustrated below.

Based on the monomials, the PIPs can be readily derived as (for a system has $N$ atoms),

$$\mathbf{G} = \hat{S} \prod_{i<j}^{N} p_{ij}^{l_{ij}}, \tag{2}$$

where $l_{ij}$ is the order of each monomial. $\hat{S}$ is the symmetrization operator, which consists of projection operators for all possible permutation operations in this system. For a system with two identical atoms, for example, $\hat{S} = (\hat{I} + \hat{X})/\sqrt{2}$ where $\hat{I}$ and $\hat{X}$ are the identity and exchange operators, respectively.

In the PIP-NN approach, the symmetrized polynomial vector, $\mathbf{G} = \{G_i\}$, is used as the input of the NN. For simplicity, only low orders ($l_{ij} \le 1$) of each monomial need to be included. For $A_3$ systems, for example, the three A–A bonds are chemically equivalent, and the following symmetrized polynomials can be used:

$$G_1 = (p_{12} + p_{23} + p_{13})/3, \tag{3a}$$

$$G_2 = (p_{12}p_{23} + p_{12}p_{13} + p_{23}p_{13})/3, \tag{3b}$$

$$G_3 = p_{12}p_{23}p_{13}. \tag{3c}$$

It is clear that the exchange any of two A atoms has no effect on the input variables, thus preserving the permutation symmetry. Also note that the orders of the three PIPs are different, rendering them linearly independent. Similarly, for $A_2B$ systems, the permutation symmetry of the two A atoms, denoted 1 and 2, can be adapted using the following PIPs:

$$G_1 = (p_{13} + p_{23})/2, \tag{4a}$$

$$G_2 = p_{13}p_{23}, \tag{4b}$$

$$G_3 = p_{12}. \tag{4c}$$

It should be pointed out that the above choices are not unique. Indeed, one can choose other PIPs, as long as they are not linearly dependent. For simplicity, we will only use the lowest-order ones. Similar choices of low-order linearly independent PIPs can be made for larger systems.

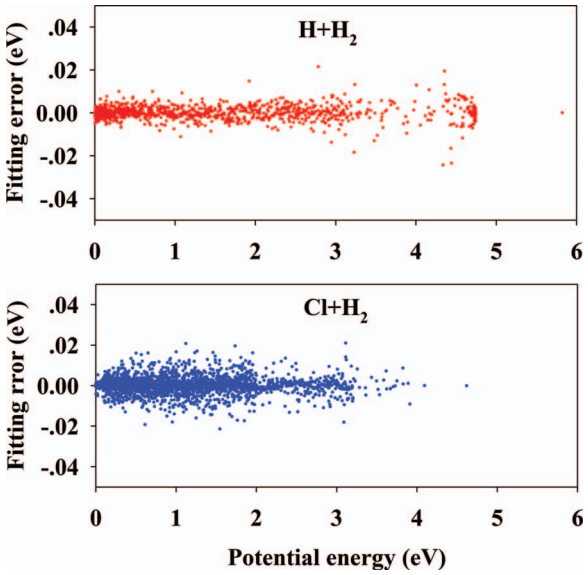


FIG. 2. Distributions of the fitting error for the H + $H_2$ and Cl + $H_2$ reactions, as a function of potential energy for the sampling points.

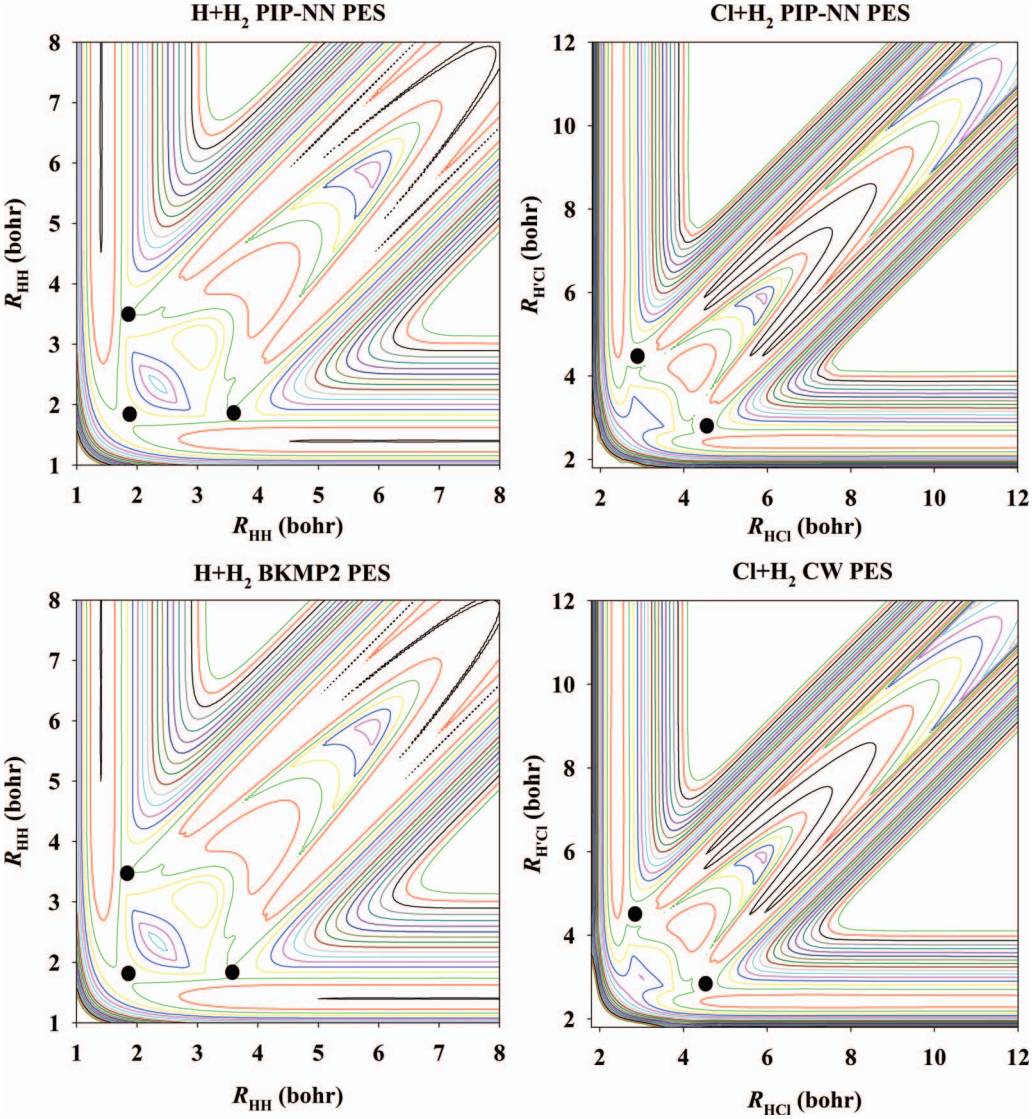


FIG. 3. Comparison of contour plots obtained from the PIP-NN and original PESs, as a function two equivalent bonds, with the enclosed angle relaxed. For the H + $H_2$ reaction, three permutationally equivalent transition states are indicated by black circles. For the Cl + $H_2$ reactions, two permutationally equivalent transition states are indicated by black circles.

The PIP-NN approach is rigorous in symmetry adaptation and has a number of advantages. First, the definition of the PIPs is simple and straightforward for molecules of any sizes and it requires no symmetrized coordinates, which can be difficult to design for polyatomic molecules with high permutation symmetry. Second, it is completely general as it requires no modification of the existing NN structure, thus permitting the use of commercial NN software.[37]

In practice, the standard forward-feed NN or multilayer perceptron is constructed with $\mathbf{G} = \{G_i\}$ being the input, as shown in Fig. 1. Without loss of generality, the value of $k$th neuron in the $i$th hidden layer can be written as[14]

$$y_k^i = f_i\left(b_k^i + \sum_{j=1}^{N_{i-1}} w_{jk}^i y_j^{i-1}\right), \quad 1 \le i \le m, \tag{5}$$

where $N_{i-1}$ is the number of neurons in the $(i-1)$th layer, which equals the number of symmetrized polynomials of the input layer when $i = 1$, $f_i$ are transfer functions for the $i$th layer, $w_{jk}^i$ are weights that connect the $j$th neurons in the $(i-1)$th layer and the $k$th neurons in the $i$th layer and $b_k^i$ are the biases of the $k$th neurons of the $i$th layer, both of which act as an adjustable offset of the transfer functions. The output energy is expressed analogously as a single neuron in the $(m+1)$ layer,

$$E = f_{m+1}\left(b_1^{m+1} + \sum_{j=1}^{N_m} w_{j1}^{m+1} y_j^m\right). \tag{6}$$

## III. EXAMPLES

To demonstrate the power of the PIP-NN fitting method, we chose to work on the global PESs for two prototypical triatomic reactive systems, namely the H + $H_2$ and Cl + $H_2$ reactions. Since the purpose here is to prove the concept, no actually *ab initio* calculations were performed. Rather, the widely used *ab initio* based

Boothroyd-Keogh-Martin-Peterson (BKMP2)[38] and Capecchi-Werner (CW)[39] PESs were employed for these two systems, respectively. The training data sets were generated in the coordinate space using the Sobol sequence,[40] which is a quasi-random sequence in which the sampled points are quasi-uniformly distributed. To this end, we first divide the dynamically relevant configuration space of the PESs into the reactant channel, interaction region, and product channel. In each region, hundreds of points were sampled in the three coordinate dimensions. For the $H + H_2$ reaction, for example, points in the reactant channel were sampled in a cubic box defined by $R_{H_1H_2} \in [0.9, 3.0]$ bohr, $R_{H_2H_3} \in [4.5, 10.0]$ bohr, and $\theta_{H_1H_2H_3} \in [0, 180^\circ]$. By symmetry these points also cover the product channel. The points in the interaction region were sampled within the box defined by $R_{H_1H_2} \in [0.9, 4.0]$ bohr, $R_{H_2H_3} \in [0.9, 4.0]$ bohr, and $\theta_{H_1H_2H_3} \in [0, 180^\circ]$. Points that contain too short H-H distances were eliminated from the data set due to their high energies. In addition, the Euclidean distances[14] between a new point and all existing points in the data set were computed and the point was discarded if any of the distances is within 0.1 bohr. In order to avoid unphysical behaviors in regions that are not dynamically relevant, a small set of points were sampled in the largest cubic box defined by $R_{H_1H_2} \in [0.9, 10.0]$ bohr, $R_{H_2H_3} \in [0.9, 10.0]$ bohr, and $\theta_{H_1H_2H_3} \in [0, 180^\circ]$. Additional points were added where the discrepancy between the fitted and original PESs is large, which improve the quality of the fit. Approximately 1000 and 2600 points were assembled for the $H + H_2$ and $Cl + H_2$ systems, respectively, and they were sufficient to generate satisfactory PESs. Note that these NN PESs are not valid beyond the largest box defined in the fitting.

The NNs include two or three hidden layers with 10 neurons in each hidden layer, leading to 161 and 271 nonlinear parameters, respectively, for the two systems. The monomial was chosen to be logarithmic functions of the internuclear distances, namely $p_{ij} = \ln(-ar_{ij})$, $a = 1$ bohr$^{-1}$. The transfer functions for the hidden layers were taken as hyperbolic tangent functions, and linear function was used for the output layer to avoid any constraint in the range of output values. Both NNs were trained using the Levenberg-Marquardt algorithm[41] and the root mean square error (RMSE), defined as

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N_{\mathrm{data}}} (E_{\mathrm{output}} - E_{\mathrm{target}})^2 / N_{\mathrm{data}}}, \qquad (7)$$

was used the measure the performance. Here $E_{\mathrm{target}}$ and $E_{\mathrm{output}}$ are the energy of input data and the fitted energy, respectively. The fitting was terminated either when the RMSE satisfies the convergence criteria ($10^{-3}$ eV) or when it stops changing with further iterations.

The overall RMSEs for the $H + H_2$ and $Cl + H_2$ systems are 3.6 and 4.2 meV, respectively. The distributions of the fitting error, which is defined by the difference between the fitted and original potential energies at the sampling points, are shown in Fig. 2. As can be seen, the randomly selected points are distributed in a wide energy range, but more concentrated at low energies. This sampling strategy covers the dynamically relevant configuration space well and thus ideally suited for dynamical calculations. As shown by the figure, the fitting error is typically within the ±0.02 eV boundary for both systems. More interestingly, the fitting error does not increase dramatically with the energy.

Figure 3 illustrates the quality of the fitting for both PESs. The energy contours of the PESs for the $H + H_2$ reaction are plotted as a function of the approaching and separating H–H bonds with the enclosed angle relaxed. It is clear that the NN PES reproduces the original BKMP2 PES very well. Because the permutation symmetry is maintained by construction, the three equivalent transition states can be found in the same positions in the NN PES as those in the original PES. In addition, the contour plots show that there are no unphysical deep wells/peaks in the fitted PES. Similar behaviors can be found from the contours of the $Cl + H_2$ PESs, which are plotted as a function of the two equivalent H–Cl distances, with the H–Cl–H angle relaxed. The two permutationally identical transition states are also indicated in the PESs. It is important to point out that these identical transition states, as well as any permutationally equivalent geometries yield exactly the same energies on the NN PESs. In other words, the energy difference only exists within the machine roundoff error.

To further illustrate the quality of the NN PESs, the $J = 0$ total reaction probabilities were computed on the NN PESs and compared with those obtained with the original PESs.

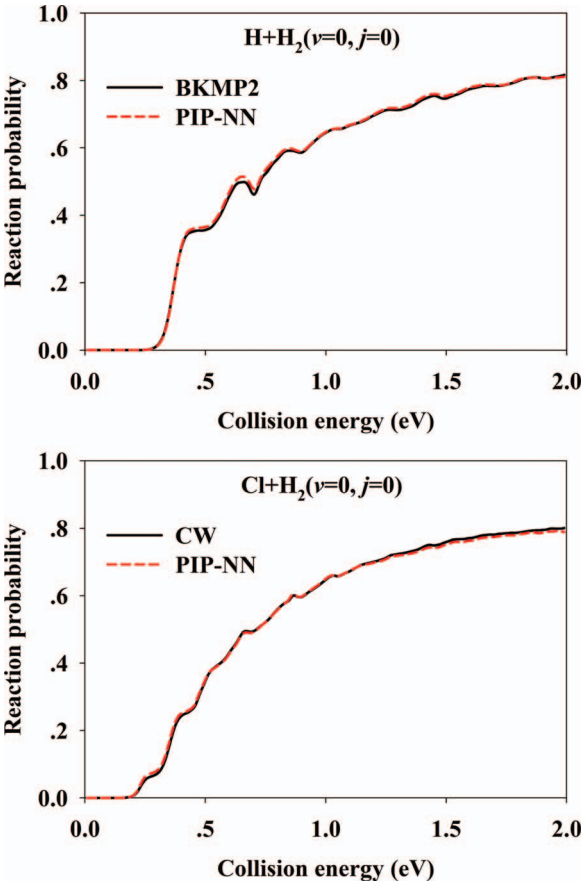


FIG. 4. Comparison of the $J = 0$ total reaction probabilities obtained from the PIP-NN and original PESs for the $H + H_2$ ($v = 0, j = 0$) and $Cl + H_2$ ($v = 0, j = 0$) reactions, respectively.

These quantum scattering calculations were performed using the Chebyshev real wave packet (CRWP) method.[42] Since the CRWP method have been described in detail in Refs. 43–46, no details is given here. The parameters used in quantum scattering calculations have been summarized in our recent work,[47] thus will not be repeated here either. As shown in Fig. 4, the original and NN PESs yield almost identical reaction probabilities, thus validating the accuracy of the NN PESs for both systems.

## IV. SUMMARY

NNs have emerged as an effective and robust tool for fitting PESs of molecular systems. However, the adaptation of permutation symmetry still represents an unsolved problem in NN-based PES fitting. In this work, we proposed a simple, rigorous, general, and efficient procedure to take advantage of permutation symmetry in NN fitting of PESs. This so-called PIP-NN approach is based on the premise that permutation symmetry can be readily adapted in the input of NNs using low-order PIPs based on monomials of internuclear distances. This approach requires neither definition of symmetrized coordinates, which can be difficult to design, nor modification of the standard NN toolboxes. As a result, it can be readily implemented in standard NN methods. Its accuracy and efficiency are demonstrated by fitting the PESs for two prototypical atom-diatom reactions.

## ACKNOWLEDGMENTS

This work was supported by the Department of Energy (DE-FG02-05ER15694). H.G. thanks Sergei Manzhos for discussions of neural network methods.

[1] J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, and A. J. C. Varandas, *Molecular Potential Energy Functions* (Wiley, Chichester, 1984).
[2] N. Sathyamurthy and L. M. Raff, J. Chem. Phys. **63**, 464 (1975).
[3] C. Xu, D. Xie, D. H. Zhang, S. Y. Lin, and H. Guo, J. Chem. Phys. **122**, 244305 (2005).
[4] A. J. C. Varandas, Adv. Chem. Phys. **74**, 255 (1988).
[5] T. Hollebeek, T.-S. Ho, and H. Rabitz, Annu. Rev. Phys. Chem. **50**, 537 (1999).
[6] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[7] M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).
[8] B. J. Braams and J. M. Bowman, Int. Rev. Phys. Chem. **28**, 577 (2009).
[9] J. M. Bowman, G. Czakó, and B. Fu, Phys. Chem. Chem. Phys. **13**, 8094 (2011).
[10] R. Dawes, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **128**, 084107 (2008).
[11] R. Dawes, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Phys. Chem. A **113**, 4709 (2009).
[12] C. M. Handley and P. L. A. Popelier, J. Phys. Chem. A **114**, 3371 (2010).
[13] J. Behler, Phys. Chem. Chem. Phys. **13**, 17930 (2011).
[14] L. M. Raff, R. Komanduri, M. Hagan, and S. T. S. Bukkapatnam, *Neural Networks in Chemical Reaction Dynamics* (Oxford University Press, Oxford, 2012).
[15] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, J. Chem. Phys. **103**, 4129 (1995).
[16] D. F. R. Brown, M. N. Gibbs, and D. C. Clary, J. Chem. Phys. **105**, 7597 (1996).
[17] H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, J. Phys. Chem. A **102**, 4596 (1998).
[18] F. V. Prudente, P. H. Acioli, and J. J. Soares Neto, J. Chem. Phys. **109**, 8801 (1998).
[19] L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, J. Chem. Phys. **122**, 084104 (2005).
[20] S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 084109 (2006).
[21] S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 194105 (2006).
[22] S. Manzhos, X. G. Wang, R. Dawes, and T. Carrington, J. Phys. Chem. A **110**, 5295 (2006).
[23] S. Lorenz, M. Scheffler, and A. Gross, Phys. Rev. B **73**, 115431 (2006).
[24] J. Ludwig and D. G. Vlachos, J. Chem. Phys. **127**, 154716 (2007).
[25] J. Behler, S. Lorenz, and K. Reuter, J. Chem. Phys. **127**, 014705 (2007).
[26] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[27] M. Malshe, L. M. Raff, M. G. Rockley, M. Hagan, P. M. Agrawal, and R. Komanduri, J. Chem. Phys. **127**, 134105 (2007).
[28] M. Malshe, R. Narulkar, L. M. Raff, M. Hagan, S. Bukkapatnam, P. M. Agrawal, and R. Komanduri, J. Chem. Phys. **130**, 184102 (2009).
[29] M. H. Le, S. Huynh, and L. M. Raff, J. Chem. Phys. **131**, 014107 (2009).
[30] H. M. Le and L. M. Raff, J. Phys. Chem. A **114**, 45 (2010).
[31] H. T. T. Nguyen and H. M. Le, J. Phys. Chem. A **116**, 4629 (2012).
[32] J. Chen, X. Xu, and D. H. Zhang, J. Chem. Phys. **138**, 154301 (2013).
[33] J. Chen, X. Xu, X. Xu, and D. H. Zhang, J. Chem. Phys. **138**, 221104 (2013).
[34] S. Haykin, *Neural Network and Learning Machines* (Pearson, New Jersey, 2009).
[35] B. G. Sumpter, C. Getino, and D. W. Noid, Annu. Rev. Phys. Chem. **45**, 439 (1994).
[36] Z. Xie and J. M. Bowman, J. Chem. Theory Comput. **6**, 26 (2010).
[37] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox™ 7 User's Guide* (The MathWorks, Inc., Natick, MA, 2010).
[38] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **104**, 7139 (1996).
[39] G. Capecchi and H. J. Werner, Phys. Chem. Chem. Phys. **6**, 4975 (2004).
[40] I. M. Sobol, USSR Comput. Math. Math. Phys. **7**, 86 (1967).
[41] M. T. Hagan and M. B. Menhaj, IEEE Trans. Neural Networks **5**, 989 (1994).
[42] H. Guo, Int. Rev. Phys. Chem. **31**, 1 (2012).
[43] S. Y. Lin and H. Guo, Phys. Rev. A **74**, 022703 (2006).
[44] S. Y. Lin and H. Guo, J. Chem. Phys. **119**, 11602 (2003).
[45] Z. Sun, S.-Y. Lee, H. Guo, and D. H. Zhang, J. Chem. Phys. **130**, 174102 (2009).
[46] J. Ma, S. Y. Lin, H. Guo, Z. Sun, D. H. Zhang, and D. Xie, J. Chem. Phys. **133**, 054302 (2010).
[47] B. Jiang and H. Guo, J. Chem. Phys. **138**, 234104 (2013).