# Representing Global Reactive Potential Energy Surfaces Using Gaussian Processes

Brian Kolb,[†,‡] Paul Marshall,[†,$] Bin Zhao,[†,#] Bin Jiang,[§] and Hua Guo[*,†]

[†]Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, New Mexico 87131, United States
[‡]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[§]Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, China

ABSTRACT: Representation of multidimensional global potential energy surfaces suitable for spectral and dynamical calculations from high-level ab initio calculations remains a challenge. Here, we present a detailed study on constructing potential energy surfaces using a machine learning method, namely, Gaussian process regression. Tests for the $^3A''$ state of SH_2, which facilitates the $SH + H \leftrightarrow S(^3P) + H_2$ abstraction reaction and the $SH + H' \leftrightarrow SH' + H$ exchange reaction, suggest that the Gaussian process is capable of providing a reasonable potential energy surface with a small number ($\sim 1 \times 10^2$) of ab initio points, but it needs substantially more points ($\sim 1 \times 10^3$) to converge reaction probabilities. The implications of these observations for construction of potential energy surfaces are discussed.
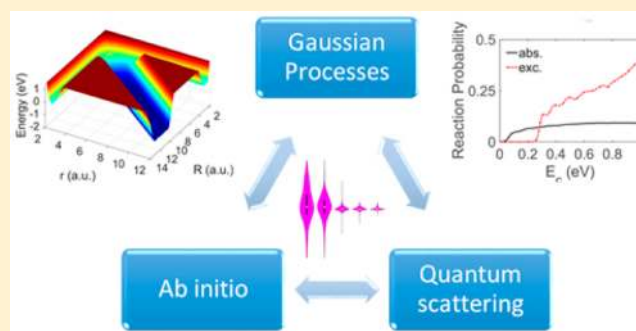
I. INTRODUCTION

The Born−Oppenheimer potential energy surface (PES), namely, the adiabatic potential energy of a molecular system as a function of its nuclear coordinates, plays a fundamentally important role in physical chemistry.[1,2] It governs the nuclear motion and gives rise to spectral and dynamical properties of the system, which eventually contribute to its thermodynamics and kinetics.[3] Recent advances in electronic structure theory have enabled accurate calculations of adiabatic potential energies for small systems at any arbitrary molecular configuration.[4] However, faithfully representing the global PES from these ab initio points remains a challenge. This is particularly true for a reactive system in which a global PES in the entire configuration space between reactants and products needs to be characterized accurately for quantum or classical scattering calculations.[5] While traditional interpolation methods are effective in small systems,[6−9] recent attention has been focused on more efficient means to represent high-dimensional global PESs in larger systems.[10−12] Some of these efforts have taken advantage of emerging techniques in machine learning.[13] For example, feed-forward artificial neural networks (NNs) have been very successful in providing an extremely flexible and relatively straightforward method for fitting PESs to a large number of ab initio points.[14−18] In our recent work, for example, a 60-dimensional PES for molecular scattering from a metal surface was fit using a NN method.[19]

Another popular machine learning strategy is the Gaussian process (GP) regression.[20] Instead of an analytical NN function with parameters optimized by training with data points, GP provides a statistical estimate of the potential energy in an unknown geometry based on the known ab initio points. This is done by means of Bayesian inference in the functional space, and it thus has no fitting parameters. The GP PESs are smooth and differentiable, and in general they represent the known ab initio points accurately. In addition, the quality of the estimations improves systematically as more points are added. Perhaps most importantly, it has been argued that GP can provide a good estimate for an N-dimensional system with only $\sim 10N$ well-selected points,[21] and there is some evidence in support of this claim in PES fitting.[22] Therefore, it could in principle offer an efficient way to represent high-dimensional PESs.

While much work has so far been done using GP to represent high-dimensional PESs for systems near equilibrium,[23−26] there have been few attempts to tackle reactive systems.[22] A chemical reaction by definition is far from equilibrium, as reactive trajectories necessarily access high-energy regions of the PES between the reactant and product asymptotes. As a result, it is much more challenging to accurately represent the global reactive PES across a large energy range and configuration space than to map out the PES near equilibrium configurations near the potential wells. In this publication, we examine the GP representation of the lowest triplet PES for SH_2 based on varying numbers of ab initio points in an attempt to understand the behavior of GP in predicting potential energies in configurations that are not included in the ab initio data set. In addition to the commonly used error assessments, we use quantum scattering to quantify the quality of the reactive PES.

In other words, special attention is paid to the convergence of dynamical results with respect to the number of points used to build GP PESs. This is important, as the prediction errors alone do not necessarily provide definitive information on the quality of the PES for scattering calculations, because reactive scattering depends exquisitely on the details of the PES. Such knowledge has important implications in future applications of the GP method in general. This publication is organized as follows. Section II presents the GP method and its implementation, along with details of the ab initio and quantum scattering calculations. The results are presented and discussed in Section III. Finally, the conclusions and prospect of this approach are given in Section IV.

II. METHODS

A. Ab Initio Calculations. Energies of SH_2 configurations were obtained via explicitly correlated unrestricted coupled cluster theory with single, double, and perturbative triple excitations (UCCSD(T)-F12b),[27] based on spin-restricted open-shell Hartree–Fock (ROHF) wave functions, using the aug-cc-pVTZ basis set.[28,29] All calculations were performed with the Molpro 2010 program,[30] and the wave functions were constrained to $^3A''$ symmetry. A frozen-core of the sulfur 1s, 2s, and 2p orbitals was excluded from the correlation treatment.

B. Gaussian Processes. Gaussian processes (GPs) are kernel-based supervised statistical learning models, which have been widely applied to regression and classification problems in machine learning.[20] More recently, they have been used to solve physical chemistry problems, for example, constructing PESs[22–26] and modeling collision dynamics.[31,32] Rather than giving any specific functional form and then optimizing the corresponding coefficients, a GP represents a collection of random observations, any finite number of which have a joint Gaussian or normal distribution. The Gaussian distribution is defined by its mean function and the covariance function. Supposing we have n variables $\mathbf{x} = \{x_1, ..., x_n\}$, where x_i can be a vector representing internal coordinates of a molecular system, and corresponding n observations $\mathbf{y} = \{y_1, ..., y_n\}$, say the potential energies, this joint multivariate Gaussian distribution N can be expressed as

$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & ... & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & & \\ \vdots & & \ddots & \vdots \\ k(x_n, x_1) & & ... & k(x_n, x_n) \end{bmatrix} \right)
$$

(1)

or in the vector-matrix form as

$$
\mathbf{y} \sim N(\mathbf{0}, \mathbf{K}(\mathbf{x}, \mathbf{x}))
$$

(2)

where the mean function is taken to be zero for simplicity, and the covariance or kernel function $k(x_i, x_j)$ represents the similarity of two variables x_i and x_j, that is, the covariance of these two; and $\mathbf{K}(\mathbf{x},\mathbf{x})$ is the covariance matrix for the entire data set. Given the property of GP, an unknown point to be predicted (x_*, y_*) would also follow this prior distribution, so we have

$$
\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim N \left(\mathbf{0}, \begin{array}{cc} \mathbf{K}(\mathbf{x}, \mathbf{x}) & \mathbf{K}_*^T(x_*, \mathbf{x}) \\ \mathbf{K}_*(\mathbf{x}, x_*) & \mathbf{K}_{**} \end{array} \right)
$$

(3)

where $\mathbf{K}_{**} = k(x_*, x_*)$, $\mathbf{K}_*$ is the vector that consists of the covariance between x_* and all existing data, and $\mathbf{K}_*^T$ is its transpose. Interestingly, the conditional distribution of y_* for the given $\mathbf{y}$ is itself Gaussian-distributed.[20]

$$
P(y_* | \mathbf{y}) \sim N(\mathbf{K}_* \mathbf{K}^{-1} \mathbf{y}, \mathbf{K}_{**} - \mathbf{K}_* \mathbf{K}^{-1} \mathbf{K}_*^T)
$$

(4)

As a result, a GP model does not yield a single prediction of y_*, like NNs do, but its entire distribution in which the mean of y_* is

$$
\bar{y}_* = \mathbf{K}_* \mathbf{K}^{-1} \mathbf{y}
$$

(5)

and its variance is

$$
\Delta y_* = \mathbf{K}_{**} - \mathbf{K}_* \mathbf{K}^{-1} \mathbf{K}_*^T
$$

(6)

The statistical mean in eq 5 is taken as the prediction with uncertainty given in eq 6. In practice, the data may be noisy, for example, $y = f(x) + \varepsilon$, where error is also Gaussian-distributed, that is, $\varepsilon \approx N(0, \delta_n^2)$ with δ_n being the noise variance. The noise terms are then taken into account in the diagonal elements of the covariance matrix, that is, $\mathbf{K}_{ii} = k(x_i, x_i) + \delta_n^2$.

GP is a probabilistic model based on Bayesian analysis, which relies explicitly on the correlation between the data themselves and is thus referred as a "non-parametric" model.[20] The GP model is closely related to the reproducing kernel Hilbert space (RKHS) and spline models,[20] which are better known as interpolation methods in the reaction dynamics community.[7] It is worthwhile to point out that the original data are required in such a non-parametric model when making predictions, while in parametric models, such as NNs and polynomial expansions, only the fitting parameters are required once they are determined. However, note that the kernel or covariance function contains hyper-parameters that are optimized in GP. The choice of the covariance function is quite flexible. In this work, the so-called neural network covariance function,[33] given by

$$
k(x_i, x_j) = \sigma_f^2 \sin^{-1} \left[\frac{x_i^T x_j}{\sqrt{(\lambda^2 - x_i^2)(\lambda^2 - x_j^2)}} \right]
$$

(7)

was used. σ_f is the standard deviation, and λ is a characteristic correlation length, which can be dimension-specific. Several other covariance functions were tested, including squared-exponential and various members of the Matérn family. Different choices did not change the qualitative results of the present work.

One must also choose a likelihood function and a prior mean. These combined choices result in a small set of hyper-parameters that must be determined, most commonly by maximizing the log marginal likelihood function, given by

$$
\log P(\mathbf{y}|\mathbf{x}, \theta) = -\frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{1}{2} \log|\mathbf{K}| - \frac{n}{2} \log 2\pi
$$

(8)

using any multivariate optimization algorithm, such as quasi-Newton methods. Once the hyper-parameters are determined, eqs 5 and 6 can yield the prediction y_* and its variance Δy_* at a new point x_*.

The determination of the hyper-parameters requires the inversion of the covariance matrix, which scales as the cube of the number of training examples. As a result, it becomes less efficient when the number of points increases beyond $\sim 1 \times 10^4$. In addition, the estimate of potential energy scales as n^2

due to the matrix multiplications. These steep scaling laws could limit the applicability of this method in representing PESs unless only a small number of points is required.

In this work, the GP modeling was performed in MATLAB, using the Gaussian Process for Machine Learning (GPML) code of Rasmussen and Nickisch (version 3.6).[20] For the SH_2 system, optimal results were obtained using exact inference with a Gaussian likelihood (with the variance σ_n as the lone hyper-parameter), a constant mean function (with hyper-parameter c), and the covariance function given in eq 7 (with hyper-parameters σ_f and λ), resulting in a total of four hyper-parameters. As stated above, values for these parameters were selected by optimizing the log marginal likelihood using MATLAB's constrained optimization routines.

The triatomic configuration was given to the GP as a set of three interatomic distances. Since the two hydrogen atoms are indistinguishable, leading to a twofold exchange symmetry between them, the S–H distances were presented to the model with the shorter one always first. The shorter S–H distance was allowed to range between 0.9 and 5 Å, and the H–H distance varied between 0.4 and 4 Å. Structures with energies more than 2.5 eV above the asymptotic H_2 + S were not used in the fitting. This configuration space was partitioned into three regions, H_2+S, the interacting region, and SH+H. Training points were selected from each region via the Latin hypercube sampling (LHS) approach as implemented in the "lhs" R package.[34] Problem regions (those with significant oscillations or unphysical wells) were addressed by adding training points in batches of a few hundred, which are selected using the same LHS approach. We note in passing that the LHS method might not be the most efficient sampling method, but it is sufficient for the current problem, particularly when individual regions defined above are used. The final fits were tested against a test set consisting of 424 points selected within the relevant region using the same approach as the training data.

C. Quantum Scattering. Quantum reactive scattering calculations are performed to test the accuracy of the GP PESs obtained with different numbers of ab initio points. Such quantum dynamical calculations provide a stringent test of the PESs. The total reaction probabilities for zero total angular momentum are computed for both the abstraction and exchange channels, using a standard time-dependent wave packet method.[35] Specifically, a Gaussian wave packet defined in the reactant Jacobi coordinates is launched from the SH + H asymptote and propagated using the split-operator time propagator.[36] The Hamiltonian is given in the corresponding Jacobi coordinates, and an L-shaped grid[37] was used. The total reaction probabilities into both the abstraction and exchange channels are obtained using a flux method with a dividing surface placed in the product channels. The wave packet is absorbed in both product channels. The parameters used in the calculation are listed in Table I.

III. RESULTS

The hydrogen abstraction reaction

$$SH + H \rightarrow S(^3P) + H_2 \tag{R1}$$
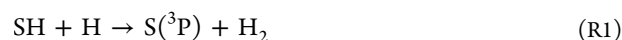
has been extensively studied and thus serves as a proto-type.[38−42] Reaction R1 proceeds on both the $^3A'$ and $^3A''$ state PESs, characterized by collinear S–H–H transition states that are quite alike. The corresponding PESs also facilitate the exchange reaction:

Table I. Parameters Used[a] in the Quantum Dynamics Calculations in an L-Shaped Scheme

grid/basis range and size

$R \in [1.5,21.0]$, $N_R^1 = 228$, $N_R^2 = 144$

$r \in [2.0,12.0]$, $N_r^1 = 600$, $N_r^2 = 90$

$j \in [0,100]$

initial wave packet: $\exp\{-(R - R_0)^2/(2\Delta_R^2)\}\exp\{-ik_0R\}$

$R_0 = 13.5$

$\Delta_R = 0.4$

$k_0 = \sqrt{2E_0\mu_R}$ with $E_0 = 0.5$ eV

absorbing potentials: $-C_x\left(\frac{x - x_e}{x_s - x_e}\right)^{n_x}$

$C_R = 0.018$, $n_R = 2$, $R_s = 17.0$

$C_r = 0.018$, $n_r = 2$, $R_s = 8.0$

[a] All units are given in atomic units unless stated otherwise.

$$SH + H' \rightarrow SH' + H \tag{R2}$$

Reaction R1 is exothermic with a small barrier, while reaction R2 is thermoneutral with a much higher barrier. There have been several previous studies of the PESs.[38−40] The PESs by Lv et al.[39,40] were obtained at the multireference configuration interaction (MRCI) level of theory, which is more accurate than the earlier PES by Maiti et al.[38] For this system, which is dominated by a single reference, the CCSD(T) PES developed here is expected to be more reliable because of the higher-order excitations included, as demonstrated recently in some recent studies of reactive PESs.[43]

Tests with geometries and energies obtained via UCCSD-(T)-F12b/aug-cc-pVTZ theory yield a barrier and overall energy change of 8.2 and −92.4 kJ mol^{-1} for R1. This is in good agreement with the MRCI values (8.4 and −91.4 kJ mol^{-1}) of Lv et al.[39] Use of the larger aug-cc-pVQZ basis set at the same geometry raises the barrier only by 0.3 kJ mol^{-1} and raises the reaction energy by 0.1 kJ mol^{-1}. These negligible changes suggest that convergence with respect to basis set size has been effectively reached at the triple-ζ level. As a result, the ab initio points were computed at the UCCSD(T)-F12b/aug-cc-pVTZ level. For the diatomic species H_2 and SH, this level of theory yields bond lengths (r_e) of 0.742 and 1.343 Å, respectively, in excellent accord with experimental values of 0.741 and 1.341 Å.[44] Similarly, the computed harmonic frequencies (ω_e) of 4401 and 2696 cm^{-1}, respectively, lie within 1 cm^{-1} of the experimentally derived values.[44] Inclusion of unscaled zero-point vibrational energy yields a 0 K enthalpy barrier and reaction enthalpy (ΔH_0) of 8.05 and −82.23 kJ mol^{-1}, respectively. As a check on the accuracy of the theory employed, we note that the latter quantity compares well with the experimental value of −81.4 ± 1.2 kJ mol^{-1}.[45,46]

A total of 3741 points were computed at the UCCSD(T)-F12b/AVTZ level of theory. These points were chosen to cover the entire configuration space relevant to R1 and R2. Only points with energies up to 148 kJ mol^{-1} above the SH + H asymptote were chosen. Five PESs were determined using the GP method described in Section II, with 250, 500, 1000, 2000, and all 3741 points. These points were selected randomly from the ab initio data set. The fitting and prediction root-mean-square errors (RMSEs) of these PESs are given in Table II, and error distributions are shown in Figure 1. To compare the PESs, the contour plots of four PESs are shown in Figure 2, in which all the four versions are very close in the dynamically accessible region of the PES.

Table II. Fitting and Prediction RMSEs of the Four PESs in Predicting the Entire ab Initio Data Set

| N	E_{fit} (meV)	$E_{\text{prediction}}$ (meV)
250	0.8	26.7
500	2.0	10.3
1000	3.0	4.3
2000	1.2	1.6
all (3741)	2.1	0.88
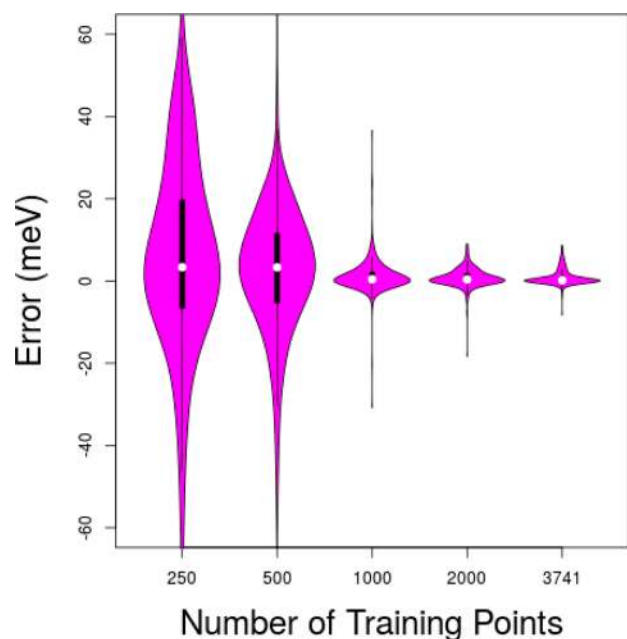
Figure 1. Violin plots of the error distribution for each of the five fits performed here as a function of the number of training points in the fit. For each size of training set the plot shows the median value (○), the range containing 50% of the data (the dark vertical black bar), and a kernel density plot showing the full distribution of the errors (the shaded area). For example, for the 500 training point set the median error is very close to 0 meV, 50% of the errors lie within ±10 meV, and almost all the remaining errors lie within ±40 meV, though there are outliers beyond this range. Note that, for visibility of the more accurate sets, the y-axis has been made smaller than would be necessary to show the entire range of errors for the 250 and 500 training point sets. It is apparent that the fits can become quite accurate, but more than ∼500 points are required for quantitative accuracy.

In Figure 3, the $J = 0$ reaction probabilities for both the abstraction and exchange channels are displayed. The abstraction channel has a very small reaction threshold, thanks to the relatively low barrier. However, its reactivity is much lower than that of the exchange channel, which has a higher threshold because of its higher barrier. For the PES with 250 points, the reaction probabilities in both channels deviate significantly from the converged values. This is consistent with the relatively large prediction errors of this PES shown in Table II and in Figure 1. When the number of data points is increased to 500, there are significant improvements for both channels, but the exchange probability still contains substantial errors. In all other three cases, the reaction probabilities of both channels are reasonably well converged. For the PES with 2000 points, the results are almost identical to the one trained with all the points, indicating that 2000 points are sufficient to provide a quantitatively converged PES for these reactions.
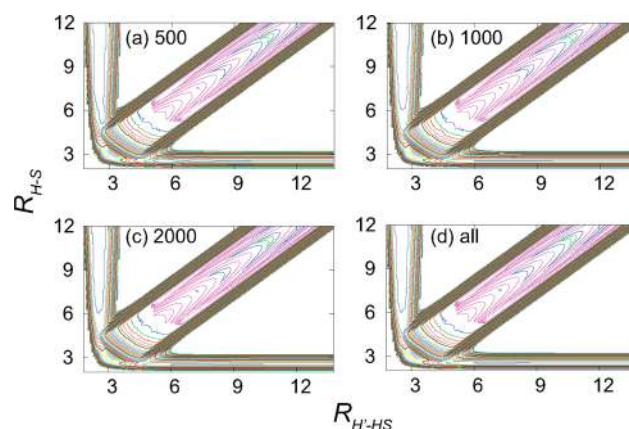
Figure 2. Contour plots of the four versions of PESs in the two radial Jacobi coordinates (in bohr) with (a) 500, (b) 1000, (c) 2000, and (d) all data points. The Jacobi angle is optimized.
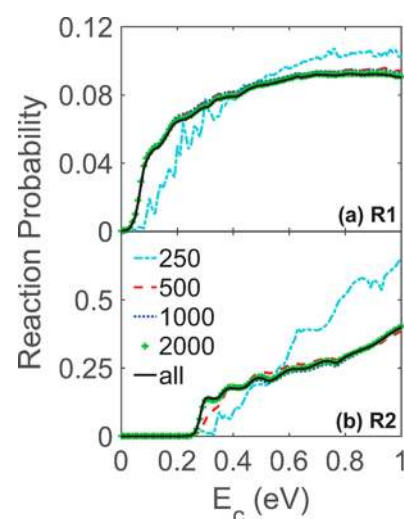
Figure 3. $J = 0$ reaction probabilities for both the abstraction (R1) and exchange (R2) channels on five PESs trained with 250, 500, 1000, 2000, and all 3741 points.

An interesting observation of these results is that the fitting errors, namely, RMSEs for the points included in the GP representation, are not exactly zero, albeit quite small. This suggests that the GP means do not always pass through the ab initio points, although the deviations are quite small. This is due to two factors. One is the choice of the covariance function, which is not as flexible as some other such functions. The other factor is the small noise included in the GP modeling, which improves the numerical stability. In addition, the prediction errors are all small, with perhaps the exception of the 250-point set. However, the calculated quantum reaction probabilities clearly demonstrate that small prediction errors do not necessarily lead to convergence in observables. This is an important point, as the previous work of Cui and Krems[22] has demonstrated small prediction errors in representing PESs but did not provide evidence for convergence of measurable quantities for dynamics.

These results also indicated that that the $10N$ rule for representing PESs depends on the level of accuracy required by the physical problem. For this three-dimensional reactive system, it is apparent that 30 points cannot produce a sufficiently accurate PES. The minimal number in this particular

system seems to be 500. This itself is already quite impressive, as thousands of points are typically needed for an NN fit,[18] and over 6000 points were used by Lv et al.[39] to construct a PES for the same system using many-body expansion with an RMSE of ~19 meV. It should probably be pointed out that the topography of the reactive PES is intrinsically more complex than that near the equilibrium geometry. Thus, it is possible that high-dimensional near-equilibrium PESs can still be represented by GP with a small number of points. For reactive systems, it is important to test more systems to establish the applicability of the GP approach and to see if the scaling is linear with respect to dimensionality.

IV. CONCLUSIONS

In this work, we have explored the possible GP representation of reactive PESs in a triatomic system. The results are promising, but more work is certainly needed to establish its applicability, especially in higher dimensions. The results presented in this work suggest that, at least for the current system, a smaller number of points is needed to produce a reasonably good PES, but significantly more points are required to produce quantitatively converged reaction probabilities. This is a common behavior of Bayesian inference-based methods. This observation underscores the difficulties associated with representing global PESs with even accuracy in all configurations, including those far from the potential wells. More studies in higher dimensionality are needed to verify the scalability of this method. In addition, an efficient and rigorous symmetry adaptation scheme is needed to enforce the permutation invariance of the PESs.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: hguo@unm.edu.

ORCID ⓘ
Bin Jiang: 0000-0003-2696-5436
Hua Guo: 0000-0001-9901-053X

Present Address
#Theoretische Chemie, Fakultät für Chemie, Universität Bielefeld, Universitätsstr. 25, D-33615 Bielefeld, Germany.

Notes
The authors declare no competing financial interest.
$On sabbatical leave from Department of Chemistry, University of North Texas, Denton, Texas 76203−5070.

■ REFERENCES

(1) Murrell, J. N.; Carter, S.; Farantos, S. C.; Huxley, P.; Varandas, A. J. C. *Molecular Potential Energy Functions*; Wiley: Chichester, England, 1984.
(2) Schatz, G. C. The analytical representation of electronic potential-energy surfaces. *Rev. Mod. Phys.* **1989**, *61*, 669.
(3) Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*; Prentice Hall: Englewood Cliffs, NJ, 1989.

(4) Dawes, R.; Ndengué, S. A. Single- and multireference electronic structure calculations for constructing potential energy surfaces. *Int. Rev. Phys. Chem.* **2016**, *35*, 441−478.
(5) Zhang, D. H.; Guo, H. Recent advances in quantum dynamics of bimolecular reactions. *Annu. Rev. Phys. Chem.* **2016**, *67*, 135−158.
(6) Varandas, A. J. C. Intermolecular and intramolecular potentials: Topographical aspects, calculation, and functional representation via a double many-body expansion method. *Adv. Chem. Phys.* **1988**, *74*, 255−338.
(7) Hollebeek, T.; Ho, T.-S.; Rabitz, H. Constructing multidimensional molecular potential energy surfaces from ab initio data. *Annu. Rev. Phys. Chem.* **1999**, *50*, 537−570.
(8) Collins, M. A. Molecular potential-energy surfaces for chemical reaction dynamics. *Theor. Chem. Acc.* **2002**, *108*, 313−324.
(9) Majumder, M.; Ndengué, A. S.; Dawes, R. Automated construction of potential energy surfaces. *Mol. Phys.* **2016**, *114*, 1−18.
(10) Braams, B. J.; Bowman, J. M. Permutationally invariant potential energy surfaces in high dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577−606.
(11) Bowman, J. M.; Czakó, G.; Fu, B. High-dimensional ab initio potential energy surfaces for reaction dynamics calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 8094−8111.
(12) Li, J.; Jiang, B.; Song, H.; Ma, J.; Zhao, B.; Dawes, R.; Guo, H. From ab initio potential energy surfaces to state-resolved reactivities: The X + H_2O ↔ HX + OH (X = F, Cl, and O(^{3}P)) reactions. *J. Phys. Chem. A* **2015**, *119*, 4667−4687.
(13) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **2016**, *145*, 170901.
(14) Handley, C. M.; Popelier, P. L. A. Potential energy surfaces fitted by artificial neural networks. *J. Phys. Chem. A* **2010**, *114*, 3371−3383.
(15) Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.
(16) Raff, L. M.; Komanduri, R.; Hagan, M.; Bukkapatnam, S. T. S. *Neural Networks in Chemical Reaction Dynamics*; Oxford University Press: Oxford, England, 2012.
(17) Manzhos, S.; Dawes, R.; Carrington, T. Neural network-based approaches for building high dimensional and quantum dynamics-friendly potential energy surfaces. *Int. J. Quantum Chem.* **2015**, *115*, 1012−1020.
(18) Jiang, B.; Li, J.; Guo, H. Potential energy surfaces from high fidelity fitting of ab initio points: The permutation invariant polynomial-neural network approach. *Int. Rev. Phys. Chem.* **2016**, *35*, 479−506.
(19) Kolb, B.; Luo, X.; Zhou, X.; Jiang, B.; Guo, H. High-dimensional atomistic neural network potentials for molecule−surface interactions: HCl scattering from Au(111). *J. Phys. Chem. Lett.* **2017**, *8*, 666−672.
(20) Rasmussen, C. E.; Williams, C. K. I. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, 2006.
(21) Loeppky, J. L.; Sacks, J.; Welch, W. J. Choosing the sample size of a computer experiment: A practical guide. *Technometrics* **2009**, *51*, 366−376.
(22) Cui, J.; Krems, R. V. Efficient non-parametric fitting of potential energy surfaces for polyatomic molecules with Gaussian processes. *J. Phys. B: At., Mol. Opt. Phys.* **2016**, *49*, 224001.
(23) Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. A. Optimal construction of a fast and accurate polarisable water potential based on multipole moments trained by machine learning. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6365−6376.
(24) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.
(25) Mills, M. J. L.; Hawe, G. I.; Handley, C. M.; Popelier, P. L. A. Unified approach to multipolar polarisation and charge transfer for ions: microhydrated Na⁺. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18249−18261.
(26) Alborzpour, J. P.; Tew, D. P.; Habershon, S. Efficient and accurate evaluation of potential energy matrix elements for quantum

dynamics using Gaussian process regression. *J. Chem. Phys.* **2016**, *145*, 174112.

(27) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(28) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(29) Woon, D. E.; Dunning, T. H. Gaussian-basis sets for use in correlated molecular calculations. 3. the atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358.

(30) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242−253.

(31) Cui, J.; Krems, R. V. Gaussian process model for collision dynamics of complex molecules. *Phys. Rev. Lett.* **2015**, *115*, 073202.

(32) Cui, J.; Li, Z.; Krems, R. V. Gaussian process model for extrapolation of scattering observables for complex molecules: From benzene to benzonitrile. *J. Chem. Phys.* **2015**, *143*, 154101.

(33) Neal, R. M. *Bayesian Learning for Neural Networks*; Springer: New York, 1996.

(34) Stein, M. Large sample properties of simulations using Latin hypercube sampling. *Technometrics* **1987**, *29*, 143−151.

(35) Zhang, J. Z. H. *Theory and Application of Quantum Molecular Dynamics*; World Scientific: Singapore, 1999.

(36) Feit, M. D.; Fleck, J. A., Jr.; Steiger, A. Solution of the Schroedinger equation by a spectral method. *J. Comput. Phys.* **1982**, *47*, 412−433.

(37) Mowrey, R. C. Reactive scattering using efficient time-dependent quantum mechanical wave packet methods on an L-shaped grid. *J. Chem. Phys.* **1991**, *94*, 7098−7105.

(38) Maiti, B.; Schatz, G. C.; Lendvay, G. Importance of intersystem crossing in the S(^{3}P, ^{1}D) + H$_2$ → SH + H reaction. *J. Phys. Chem. A* **2004**, *108*, 8772−8781.

(39) Lv, S.-J.; Zhang, P.-Y.; Han, K.-L.; He, G.-Z. Exact quantum scattering study of the H + HS reaction on a new ab initio potential energy surface H$_2$S (^{3}A″). *J. Chem. Phys.* **2012**, *136*, 094308.

(40) Lv, S.-j.; Zhang, P.-y.; He, G-z. Time-dependent wave packet quantum scattering study of reaction S(^{3}P)+H$_2$→HS+H on a new ab initio potential energy surface ^{3}A′. *Chin. J. Chem. Phys.* **2012**, *25*, 291−296.

(41) Duan, Z. X.; Li, W. L.; Xu, W. W.; Lv, S. J. Quasiclassical dynamics for the H + HS abstraction and exchange reactions on the ^{3}A″ and the ^{3}A′ states. *J. Chem. Phys.* **2013**, *139*, 094307.

(42) Wu, H.; Duan, Z.-X.; Yin, S.-H.; Zhao, G.-J. State-resolved dynamics study of the H + HS reaction on the ^{3}A′ and ^{3}A″ states with time-dependent quantum wave packet method. *J. Chem. Phys.* **2016**, *145*, 124305.

(43) Nguyen, T. L.; Li, J.; Dawes, R.; Stanton, J. F.; Guo, H. Accurate determination of barrier height and kinetics for the F + H$_2$O → HF + OH reaction. *J. Phys. Chem. A* **2013**, *117*, 8864−8872.

(44) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 17b, S. Johnson, R. D., III, Ed.; http://cccbdb.nist.gov/.

(45) Ruscic, B. Updated Active Thermochemical Tables (ATcT) values based on ver. 1.110 of the Thermochemical Network, 2012; online at ATcT.anl.gov.

(46) Morley, G. P.; Lambert, I. R.; Mordaunt, D. H.; Wilson, S. H. S.; Ashfold, M. N. R.; Dixon, R. N.; Western, C. M. Translational spectroscopy of H(D) atom fragments arising from the photo-dissociation of H$_2$S(D$_2$S): a redetermination of D(S-H). *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 3865−3875.