THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

Cite This: J. Phys. Chem. Lett. 2019, 10, 5250−5258

pubs.acs.org/JPCL

# A Machine Learning Approach for Prediction of Rate Constants

Paul L. Houston,*[,†] Apurba Nandi,*[,‡] and Joel M. Bowman*[,‡]

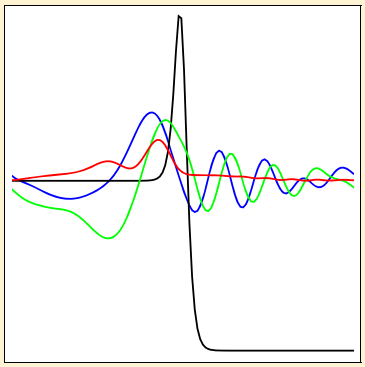
[†]Department of Chemistry & Chemical Biology, Cornell University, Ithaca, New York 14853, United States

Department of Chemistry & Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

[‡]Department of Chemistry and Cherry L. Emerson Center for Scientic Computation, Emory University, Atlanta, Georgia 30322, United States

Supporting Information

**ABSTRACT:** We report a machine learning approach to train and predict bimolecular thermal rate constants over a large temperature range. The approach uses Gaussian process (GP) regression to evaluate the difference between accurate quantum results and Eckart-corrected conventional transition state theory, mostly for collinear reactions. Training is done on a database of rate constants for 13 reaction/potential surface combinations, and testing is performed on a set of 39 reaction/potential surface combinations. Averaged over all test reactions, the GP method is within 80% of the accurate answer, whereas transition state theory (TST) is only within 330% and Eckart-corrected TST (ECK) is within 110%. In the tunneling region, GP is generally (with a few exceptions) more accurate and sometimes much more accurate. In the high-temperature recrossing region, GP is significantly more accurate than either TST or ECK, neither of which addresses the possibility of recrossing. The GP predictions for the 3D reactions $O(^3P) + H_2$, $OH + H_2$, $O(^3P) + CH_4$, and $H + CH_4$, for which accurate quantum results are available, provide further encouragement to the machine learning approach.



The calculation of the thermal rate constant for chemical reactions is one of the major objectives of theoretical and computational chemistry. The major breakthrough in the field came more than 50 years ago from transition state theory (TST). For the wide class of reactions with activation energies, the conventional version of this theory is based on properties of the relevant saddle point for the reaction, as can be found in any textbook on reaction rate theory. Early in the history of TST, it was recognized that for many reactions quantum mechanical (QM) tunneling is a major effect that is not accounted for in conventional TST. The first theories to incorporate effects of tunneling were based on 1D potentials, and the one due to Eckart[1] is of particular interest here. Tunneling corrections based on this potential[2,3] continue to be used.[4−7]

With the development of exact quantum methods to describe reaction dynamics in the 1970s, it became possible, in principle, to obtain exact rate constants for a given potential energy surface (PES). These are very computationally demanding, and with a handful of exceptions, the vast majority of such calculations were done for model collinear reactions, using realistic PESs. From these calculations (see ref 8 for a compilation) and a handful of 3D ones, notably the $H + H_2$ reaction,[9] some shortcomings of 1D tunneling models such as the Eckart model were revealed, e.g., "corner-cutting" tunneling paths, recrossing of the TS, vibrationally adiabatic effects, etc. More sophisticated approximate approaches were soon developed that could incorporate these effects, at least at some level of accuracy. These included a large-curvature, corner-cutting reaction path,[8,10] reduced-dimensionality quantum methods,[11−15] instanton,[16] VPT2-based semiclassical TST,[4,5,17−20] and more recently ring polymer molecular dynamics (RPMD)[21,22] and the ring polymer instanton method.[23] Exact direct methods to obtain the rate constant[24] have been implemented using multiconfiguration time-dependent Hartree (MCTDH) theory, at least for zero total angular momentum,[25,26] and then combined with J-shifting[13] to get the full rate constant.

These methods are indeed improvements over the Eckart correction to conventional TST, which we denote as ECK; however, none is exact. Numerous tests of them indicate a level of accuracy of about a factor of 1.5−2 for the rate constant at low temperatures. For example, for $H + H_2$, $OH + H_2$, and $H + CH_4$, this accuracy was reported for the RPMD method compared to benchmark results.[22,27] This level of accuracy can be 1−2 orders of magnitude better than that of a conventional TST calculation in the moderate and deep tunneling region and also more accurate than ECK, although generally with much smaller differences. However, these methods require substantially more information about the PES than is required for an ECK calculation, which unfortunately limits the level of ab initio theory that can be used to obtain the relevant part of the full-dimensional PES. This effort grows nonlinearly with the number of degrees of freedom in the reaction system. Also, as is well known, the accuracy of the PES is a serious issue for

Received: June 23, 2019
Accepted: August 18, 2019
Published: August 18, 2019






DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

the rate constant calculation, especially at temperatures where the ratio of the barrier height to $k_BT$ is large. In many cases, this means that the PES must be accurate to a level of subkcal/mol.

By contrast, ECK requires very little information about the PES beyond what is needed for a TST calculation (see below), and therefore, the required information can be relatively easy to obtain at the highest level of electronic structure theory. This tantalizing fact notwithstanding, ECK is certainly not state-of-the-art in accuracy among the approximate methods mentioned above. A remedy to this dilemma, based on machine learning (ML) is suggested and tested here. The goal is to determine whether ML can provide an efficient method to correct ECK and thereby efficiently improve the accuracy of rate constant calculations. It is perhaps worth noting the distinction between "learning" in ML and the learning from exact quantum calculations of collinear reactions noted above, e.g., corner-cutting reaction paths, adiabatic effects, etc. ML cannot learn these important physical effects. ML is successful when predictions on a test data set are sufficiently accurate to provide a mathematical model that can be used for novel predictions. This is the goal of the present approach. It is understood that this goal may fall short of providing new insight and interpretation. However, further analysis of the ML model may possibly uncover correlations between the results and the descriptors. This is not pursued here.

The specific approach we propose is to obtain the correction to ECK to obtain the accurate rate constant by training to a large data set of exact quantum calculations for collinear reactions. This correction, denoted by $\chi(T)$, is then represented by a nonparametric ML method dependent on a small number of descriptors. We start with the three ECK parameters, denoted in the pioneering literature[2,3,28] as $\alpha_1$, $\alpha_2$, and $u^*$, and then add others as described below. $\chi(T)$ is obtained using a database of accurate rate constants over a wide temperature range. The procedure calls for training over a subset of the database, followed by testing of the model. Here we use Gaussian Process (GP) regression. This choice was made because for the small training data set that we use (see below) GP is very easy and fast to apply. As is standard with this approach (and other ML approaches), testing is needed to verify the predictive power of the learning; this is done by comparing the predictions, in this case the rate constant, against accurate $k(T)$ data not included in the training.

The details of the approach we take are as follows. Let $k(T)$ be an exact quantum rate constant over some temperature range for a given reaction and PES, and let $k^{TST}(T)$ be the conventional TST rate constant. Typically, corrections to $k^{TST}(T)$ are written as $\kappa(T)k^{TST}(T)$, where the dimensionless $\kappa(T)$ is the transmission coefficient. It ranges from the exact coefficient, which generally is not known of course, to various approximate ones, e.g., one based on the Eckart model. In principle, ML can be used to train on the exact $\kappa(T)$ and then predict rate constants based on the training. Instead, we express $k(T)$ as

$$k(T) = [\kappa_{ECK}(T)k^{TST}(T)]\chi(T) \quad (1)$$

where $\kappa_{ECK}(T)$ is the Eckart correction to conventional TST, applicable for either symmetrical or unsymmetrical barriers,[2,3,28] and $\chi(T)$ is the correction to the Eckart correction to produce agreement with the accurate rate constant. We think that this approach is better than using ML on $\kappa(T)$ itself. The reasons are based on two facts. The first is that the calculation of $\kappa_{ECK}(T)$ requires even less information than needed to obtain $k^{TST}(T)$, and the second is that this transmission coefficient does give reasonable tunneling corrections. Thus, the expectation is that the correction to the Eckart transmission coefficient is small and hopefully easy to train on. Details of the calculation of $\kappa_{ECK}$ are given elsewhere,[2,3] and we note that in those papers $\kappa_{ECK}$ is denoted by $\Gamma^*$.

Formally, $\chi(T)$ depends on parameters, also known as descriptors, that characterize a reaction. The minimal set that we consider are the three dimensionless parameters that specify $\kappa_{ECK}$, namely[2,3]

$$\alpha_1 = V_1/\omega_{im} \quad (2)$$

$$\alpha_2 = V_2/\omega_{im} \quad (3)$$

$$u^*(T) = \omega_{im}/(0.69307T) \quad (4)$$

where $V_1$ ($V_2$) is the saddle point barrier height in $cm^{-1}$ relative to the reactants (products), omitting the zero-point energy in both cases, and where the energy of the reactants is zero. $\omega_{im}$ is the magnitude of the saddle point imaginary frequency (in $cm^{-1}$), 0.69307 is an energy conversion factor, $k_B/(hc)$, and $T$ is the temperature in Kelvin.

Additional parameters are also considered; these are the skew angle and the vibrational frequency of the diatomic molecule. The skew angle $\beta$ is given by $\beta = \tan^{-1}(m_BM/m_Am_C)$, where $M$ is the total mass.[29] $\beta$ varies from 0 to 90° in the limits as $m_B$ goes to zero and infinity, respectively. This angle is important for the dynamics of collinear reactions, as discussed extensively in the literature. Small values of $\beta$ typically are associated with significant corner-cutting (and enhanced tunneling) at low temperature and recrossing of the TS at high temperatures.[29] In reaction systems with more than three atoms, we treat the system as A + BC, where A is an atom or radical, B is an atom transferred in the reaction, and C has the mass of the remaining atoms. For the two examples here, OH + $H_2$ and X + $CH_4$, the vibrational frequency of BC is taken as the one for $H_2$ and for $CH_4$ the average of the two C−H stretch frequencies, i.e., 3010 $cm^{-1}$.

To proceed with training of $\chi$, a large database of accurate, i.e., fully QM, rate constants over a wide range of temperature is needed. Unfortunately, such a database does not exist for reactions in 3D. (This is a testament to how difficult such calculations are.) However, such a database can be created for collinear (2 mathematical dimensional) reactions using extensive tables in ref 8. The reactions selected for training are given in Table 1, along with the PESs. The boldfaced entries indicate those that were selected to be in the training set, and the asterisks indicate the two 3D reactions that were included in the training. The selection of the training reactions was based on diversity in terms of barrier height, skew angle, and reaction energetics. A more extensive set of reaction parameters was selected for predictions, again making use of the tables in ref 8. The complete list is given Table S1 of the Supporting Information (SI).

For isoergic reactions, e.g., H + $H_2$ and D + $H_2$, the necessary information for the evaluation of $\kappa_{ECK}$ is available in Table 1 of Allison and Truhlar;[8] in these cases, $\alpha_1 = \alpha_2$, and for isotopologue variations of isoergic reactions, the values of $\alpha_1$ and $\alpha_2$ are unchanged, but $\omega_{im}$ is different. For nonisoergic reactions, the data in that table were supplemented using heats of formation data from NIST.[39] With this information in hand,



DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

**Table 1. Reactions Included in This Study**[a]

| num. | rxn | PES | num. | rxn. | PES |
|---|---|---|---|---|---|
| **1** | **HBrH** | **DIM-3C*** | 22 | ClTT | SPK/GSW |
| 2 | HBrD | DIM-3C | 23 | OHH | JW |
| **3** | **HFH** | **SK** | **24** | **OHH** | **ModPolCI** |
| **4** | **DFD** | **SK** | 25 | OHH | PolC |
| 5 | HClCl | KNPRY | 26 | OHH | DIM-RMOS |
| 6 | HFF | JOT-II | 27 | OHH | SL |
| **7** | **DFF** | **JOT-II** | 28 | OHH | JWS |
| 8 | TFF | JOT-II | 29 | ClDD | SPK/GSW |
| 9 | HDH | PK2 | 30 | FHH | M5 |
| 10 | HDD | TK | **31** | **ClHH** | **G3*** |
| **11** | **DDD** | **TK** | 32 | ClHH | SPK/GSW |
| **12** | **HHH** | **LSTH** | 33 | THD | LSTH |
| **13** | **HHH** | **PK2** | 34 | OHD | JWS |
| **14** | **HHH** | **TK** | 35 | ClHD | SPK/GSW |
| 15 | ODH | JWS | **36** | **ClDCl** | **BCMR** |
| **16** | **ClDH** | **SPK/GSW** | 37 | ClDCl | PK3 |
| 17 | HHD | PK2 | **38** | **ClHCl** | **BCMR** |
| 18 | DHH | PK2 | 39 | ClHCl | PK3 |
| 19 | DHH | TK | 40 | ClHBr | GLM |
| 20 | ODD | JWS | 41 | O + H−$CH_3$ | SPES[30,31] |
| 21 | FDD | M5 | 42 | H + H−$CH_3$ | LCZXZG[32,33] |
| | | | 43 | HO + H−H | NN1[34−38] |

[a]Boldface type indicates that the reaction is included in the training set, and the * indicates that the training set reaction is the 3D version. References for surfaces not otherwise noted may be found in the original database.[8]

$\kappa_{ECK}$ was obtained numerically following the approach by Johnston and Heicklen[3] as modified by Brown.[28] (See the Computatonal Details.)

As noted above, Allison and Truhlar provide the results of "accurate" exact QM calculations, conventional TST calculations, and semiclassical calculations using improved canonical variational theory (ICVT) supplemented by least action tunneling (LAT). These results are tabulated for typically eight temperatures over a large range, typically from 200 to 1500 K. Where exact QM results are not available at some temperatures, we have followed their practice of substituting ICVT/LAT results as the "accurate" ones. In some cases, mixing these two methods gives irregularities in the $\log_{10} k$ vs $1/T$ curves. In the figures that follow, we will distinguish, when needed, between the ICVT/LAT and QM points.

From these tabulations of the accurate $k(T)$ and our calculation of $\kappa_{ECK}$ together with the tabulated $k^{TST}(T)$, the exact $\chi(T)$ values were obtained. To continue with a ML approach, we need a set of variables, also known as descriptors, for $\chi$. We considered three sets of descriptors. One is the three Eckart parameters just mentioned and given by eqs 2−4. Additional descriptors are the skew angle $\beta$, given in Table 1 of Allison and Truhlar, as well as $\omega_e$(B−C), the harmonic frequency of the B−C molecule in the reaction A + BC → AB + C, available from the NIST WebBook.[39] These descriptor sets are given in Table 2, which we discuss in detail below. A listing of the numerical values of the descriptors used in this study can be found in Table S1.

We chose GP regression as the ML method to fit the data. This method is straightforward, and for the present application, it is fast to use because the size of the training data is only 104 (see below). We note that GP has been used extensively in computational chemistry to represent high-dimensional

**Table 2. Parameter Sets Investigated**

| parameters | average RMS (all test reactions) |
|---|---|
| $\alpha_1$, $\alpha_2$, $u^*$ | 0.32 |
| $\alpha_1$, $\alpha_2$, $u^*$, $\beta$ | 0.30 |
| $\alpha_1$, $\alpha_2$, $u^*$, $\beta$, $\omega_e$(B−C) | 0.25 |

PESs.[40−42] Details of the method can be found in those references as well as the general ref 43; therefore, we omit a detailed description of the method here. In principle, the GP approach can exactly fit the training data; however, this can lead to oscillatory behavior in the predictions. Thus, a small "noise" term is typically added to the training data. With the inclusion of this noise, the GP does not exactly reproduce the training data. Thus, as is typical, we report the root-mean-square (RMS) error on the training data. As for prediction, the results of the training (the major one of which is the inverse of an $N \times N$ covariance matrix, where $N$ is the size of the training data set, which as noted already is 104) are used in a straightforward expression to obtain the prediction. Clearly, there is also an error in the predictions, and this is usually determined by testing against a known exact set of data not included in the training. Because prediction is the major objective of this work, we used a set of accurate rate constants not included in the training data set as the test data set.

Although the data set of $\chi$ is fit using GP regression, the thermal rate constant is the property of interest. Thus, we present error analyses on the rate constant, both the one from the training and the ones tested. For this purpose, we use $\log_{10}(k_{GP}/k_{acc})$ as the error metric, where $k_{GP}$ and $k_{acc}$ are the GP (fit/predicted) and accurate rate constants, respectively. Clearly, if the two are equal, the error is zero. To be clear, $k_{GP} = [\kappa_{ECK}(T)k^{TST}(T)]\chi_{GP}(T)$, where $\chi_{GP}(T)$ is the GP representation of $\chi(T)$. This error metric will also be used for comparisons of $k_{TST}$ and $k_{ECK}$ to $k_{acc}$.

Both "RMS" and "average RMS" values are reported below. These are defined as follows

$$\mathrm{RMS}_i = \left( (1/j_{max}) \sum_{j}^{j_{max}} \log_{10}(k_{GP}/k_{acc})_i^2 \right)^{1/2} \quad (5)$$

where $i$ refers to one of the reaction/surface combinations of Table 1 and $j$ is an index that runs over the temperatures at which that reaction/surface was evaluated. The value of $j_{max}$ was typically 8, and the RMS for the reaction/surface thus refers to the RMS deviation of $\log_{10}$ values of the rate constants over this group of temperatures. The average RMS for a group of reactions is given by

$$\overline{\mathrm{RMS}} = (1/i_{tot}) \sum_{i}^{i_{tot}} \mathrm{RMS}_i \quad (6)$$

where $i_{tot}$ is the number of reaction/surface combinations in the group.

The selection of the training set and the precision with which it is fit has a strong influence on the test accuracy. With GP, the precision of the fit to the training set is easy to obtain, albeit with caution not to overfit. The selection of the training set has to span a sufficient range of the relevant descriptors to encompass the values encountered in the test set. In the present case where by design the test set is larger than the training set, the selection process is very important. Some experimentation was done with the training set in order to



DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

obtain good results (see below) for the test set. The result was a training set that included symmetrical barriers, unsymmetrical barriers, and two 3D reactions. The range of $\omega_{im}$ for the training set was from 330 to 2189 $cm^{-1}$, the range of the forward barrier height was from 612 to 4102 $cm^{-1}$ (1.75−11.75 kcal/mol), and the temperature range was from 200 K for all reactions and up to at least 1000 K and in many cases up to 2400 K.

For each reaction, the RMS error, defined above, was calculated from the eight or so temperatures tabulated in Allison and Truhlar.[8] Averaged over all training and test reactions in Table 1, the RMS error was found to be 0.20. This average RMS analysis was also performed for each of the three sets of descriptors (variables), and the results for all reactions in just the test set of data are given in Table 2. As seen, the smallest average RMS is obtained for the set of five variables $\alpha_1$, $\alpha_2$, $u^*$, $\beta$, and $\omega_e$(B−C). Note that the RMS error averaged over all test results of 0.25 corresponds to a factor of 1.8 error or 80% between $k_{GP}$ and $k_{acc}$. For all of the results reported below, we used the five-parameter set of variables: $\alpha_1$, $\alpha_2$, $u^*$, $\beta$, $\omega_e$(B−C).

The GP method was differently successful in prediction of rate constants for different reaction groups. Table 3 shows the average over all reactions in a particular group of the RMS of $\log_{10}(k_{GP}/k_{acc})$ taken over the temperatures in each reaction. The number in parentheses gives the value of $k_x/k_{acc} = 10^{\overline{RMS}}$, where $\overline{RMS}$ is the average RMS for the group and where x = TST, ECK, or GP. Of course, the method is most successful in predicting the reaction rates for the training set. It is next most successful in predicting results for symmetrical barriers and then unsymmetrical barriers, which differ by 30 and 60% compared to $k_{acc}$. In general, it was worse at predicting 3D reaction results. This is not surprising given that only two 3D reactions were included in the training. In all cases, the GP prediction was much more accurate than the TST prediction, and in all cases except for the X−H−$CH_3$ reactions, it was the most accurate of the three. More discussion of the applicability of training on collinear reactions to reactions in 3D is given below. Considering all reactions or all test reactions, the GP prediction was better than either TST or ECK.

**Table 3. RMS Values of $\log_{10}(k_x/k_{acc})$ Taken over All Temperatures for Each Reaction and Averaged over Various Reaction Groups, Where x = TST, ECK, or GP[a]**

| rxn. group | no. rxn./ surf. combin. | average RMS ($k_x/k_{acc}$) x = TST | x = ECK | x = GP |
|---|---|---|---|---|
| training set | 13 | 0.51 (3.3) | 0.25 (1.8) | 0.04 (1.1) |
| symmetrical barriers | 5 | 0.49 (3.1) | 0.19 (1.6) | 0.13 (1.3) |
| unsymmetrical barriers | 17 | 0.44 (2.8) | 0.25 (1.8) | 0.20 (1.6) |
| 3D reactions | 14 | 0.93 (8.6) | 0.44 (2.8) | 0.35 (2.2) |
| X−H−$CH_3$ reactions | 3 | 0.58 (3.8) | 0.34 (2.2) | 0.37 (2.3) |
| all reactions | 52 | 0.60 (4.0) | 0.30 (2.0) | 0.20 (1.6) |
| all test reactions | 39 | 0.63 (4.3) | 0.32 (2.1) | 0.25 (1.8) |

[a]Calculations were performed using the five-parameter GP prediction.

A pictorial view of the RMS errors for rate constants obtained with TST, ECK, and GP for each reaction/surface listed in Table 1 is given in Figure 1. Each reaction/surface may be identified by the number on the abscissa, which

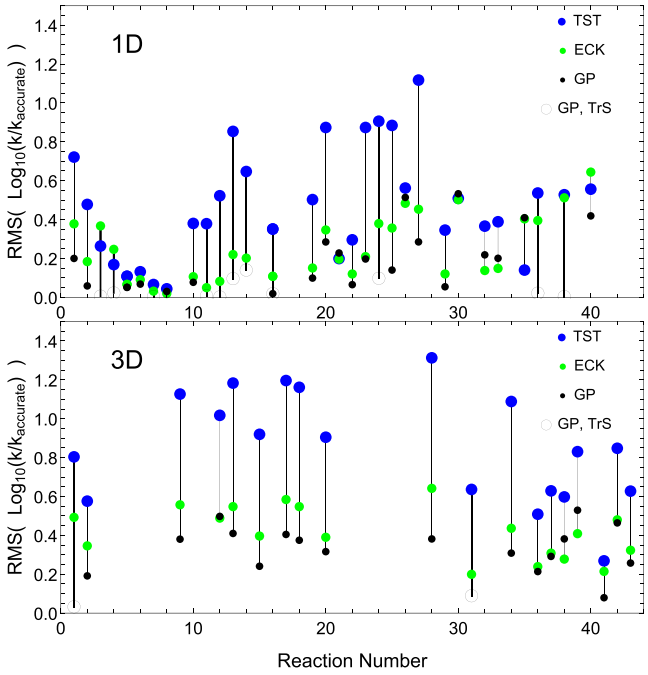



**Figure 1.** Plot of the RMS, taken over the different temperatures, of $\log_{10}(k_x) - \log_{10}(k_{acc})$ for the reactions indicated by the number in Table 1, where x is TST (blue), ECK (green), or GP (solid black for test results or open black for training set results). Results for 1D calculations are in the top panel, and those for 3D calculations are in the bottom panel. Note that a RMS of zero indicates that there is perfect correction of TST to the accurate value.

corresponds to a numbered entry in Table 1. In addition, some reactions are calculated in 1D and some in 3D and some in both. Results for 1D calculations are in the top panel, and those for the 3D calculations are in the bottom panel. The ordinate gives the RMS, taken over the temperatures, for that reaction. Comparisons are made between the TST result (blue), the ECK-corrected TST result (green), and the GP prediction (solid black for test results and open black for training results). Recall that a RMS of zero indicates perfect prediction of the accurate result. As expected, the GP error for the training reactions is usually the smallest of all three. Although the GP result is not always the best of the three estimates, it is the best in more than 81% of the total reactions and 75% of the test reactions. Notable exceptions are reactions 21 (F + $D_2$), 30 (F + $H_2$), and 35 (Cl + HD). We return to these three apparent outliers in the discussion.

Although it is important that there is good overall agreement, as discussed above, between the GP prediction and the accurate result, the detailed nature of the agreement is noteworthy. Figure 2 illustrates this point. The top panel provides the result for the collinear Cl + HCl reaction (reaction number 38) on the BCMR surface. This "heavy−light−heavy" reaction is in the training set, and one can see that the training is quite successful because the GP predictions (black circles) agree almost perfectly with the accurate quantum results (red line). The TST results are given by the blue line, and the ECK results are given by the green diamonds. In the bottom panel, we show the results for the same reaction in three dimensions. Note that this reaction is not in the training set. Although the GP predictions no longer agree perfectly with the accurate results, the qualitative changes from the TST and ECK results show that not only do the GP


DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

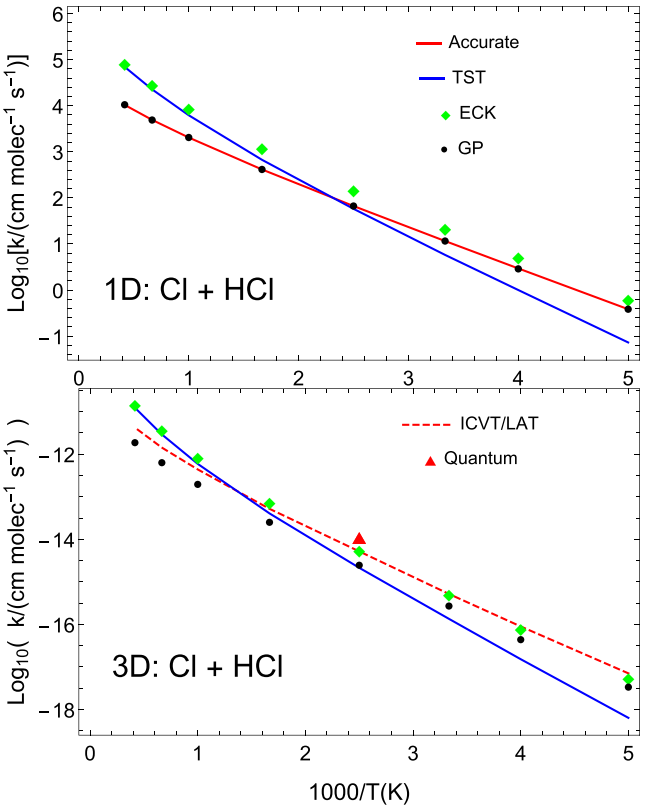



**Figure 2.** Plots of $\log_{10}(k)$ vs $1000/T$ for (top) the collinear (1D) reaction Cl + HCl → ClH + Cl and (bottom) the corresponding 3D reaction. The red curves are the "accurate" rate constants;[8] for 1D, they are exact quantum ones, and for 3D they are ICVT/LAT ones, except for the red triangle, which is an approximate quantum calculation.[44] Note the 1D ICVT/LAT results are not quantitatively accurate, and therefore, their accuracy in 3D is uncertain (see the text for more discussion). The blue curves are the TST rate constants, the green diamonds indicate the ECK results, whereas the filled black circles indicate the GP results

results correct for the tunneling at low temperatures but they also correct for recrossing at high temperatures. The ECK result corrects for the tunneling but not the recrossing, which is not addressed by the Eckart model. Recrossing is particularly important in heavy−light−heavy reactions such as this one, where the skew angle is small (here 13.6°). It should be noted that "accurate" results in the 3D reaction might well be called into question. All but the result at 1000/400 = 2.5 (red triangle) are actually ICVT/LAT results. The one exception is a quantum calculation for $T$ = 400 K based on a centrifugal-sudden distorted wave study and dating to 1990.[44] We note that for the 1D reaction the ICVT/LAT rate constants (not shown) are not particularly accurate; they are a factor of 2 too high at high temperatures and a factor of 2 too low at low temperatures. Because these ICVT/LAT points form the bulk of the dashed red "accurate" curve in the 3D reaction, it may be that the GP determinations are actually more accurate than the lower panel of Figure 2 would indicate. Finally, we note that it is remarkable how well the simple ECK rate constants are at the low temperatures; in the 3D Cl + HCl reaction, they seem to correct for most of the tunneling, although, of course, they do not address the recrossing.

Next we consider the performance of the GP predictions on three 3D reactions that are not in the training set even as collinear ones, namely, $O(^3P) + CH_4$, $H_2$ + OH, and H + $CH_4$. These reactions have been the focus of recent theoretical work using accurate quantum, RPMD, and instanton methods. They are certainly challenging reactions with significant tunneling owing to the substantial reaction barriers. On the PESs used in this work, the barriers are 13.0 kcal/mol for $O(^3P) + CH_4$, 5.4 kcal/mol for $H_2$ + OH, and 14.7 kcal/mol for H + $CH_4$.

Consider first the $O(^3P) + CH_4$ reaction. Figure 3 shows results of GP, TST,[30] RPMD,[33] and accurate quantum

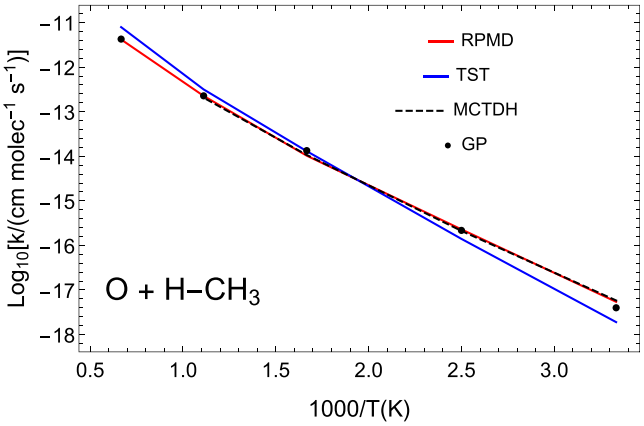



**Figure 3.** Plot of $\log_{10}(k)$ vs $1000/T$ for the 3D reaction O + H−$CH_3$ → OH + $CH_3$. The red line (RPMD) and the dashed black line (MCTDH) give the accurate rates.[30,31,33] The blue line gives the TST rates, and the filled black circles indicate the GP results.

MCTDH[31] calculations, all using the same semiempirical PES.[30] The results of the GP predictions are given by the black circles, and the TST values are given by the blue lines. The red lines and the black, dashed lines are the accurate RPMD and MCTDH results, respectively.[30,31,33] The GP method predicts quite accurately both the tunneling correction at low temperature and the recrossing correction at high ones. We defer a discussion of the success of the GP approach for this reaction until after we present the corresponding comparisons for $H_2$ + OH, H + $CH_4$, and $O(^3P) + H_2$.

Figure 4 shows results for the 3D $H_2$ + OH reaction. The red curve gives the MCTDH results,[37,38] and the dashed black

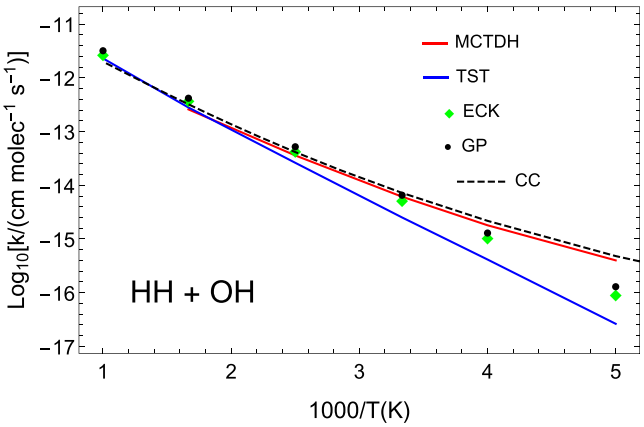



**Figure 4.** Plot of $\log_{10}(k)$ vs $1000/T$ for the 3D reaction $H_2$ + OH → H + $H_2O$. The red and black dashed curves are the accurate rate constants, from MCTDH[37,38] and coupled channel scattering calculations,[36] respectively. The blue line gives the TST rates,[35] the green diamonds give the ECK rates, and the filled black circles indicate the GP results.



DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

curve is the coupled channel (CC) scattering calculations.[36] Both use the highly accurate NN1 PES.[34] The TST results, shown by the blue line are taken from Meisner and Kästner, as corrected.[35,45] It was later realized that the correction, divided by 2 to account for the electronic degeneracy of the OH, was not correct over the entire temperature range; the correct division should be by the factor $1 + \exp(-140/(0.69307 \times T(K)))$.[36] Thus, for the TST results (blue line), we have multiplied the Meisner and Kästner results by 2 and divided by this temperature-dependent factor. As can be seen, the TST results greatly underestimate the rate constant below 500 K due to tunneling. ECK does correct this substantially, and GP produces semiquantitative agreement with the accurate MCTDH and CC results.

The rate constants for H + $CH_4$ from TST, GP, RPMD, and MCTDH calculations are shown in Figure S4. Although the GP results are much more accurate than TST in the tunneling region, they are less accurate than those for the O + $CH_4$ compared to the corresponding RPMD and MCTDH ones, which are in good agreement with each other. We discuss these results below.

The final 3D comparison is for $O(^3P)$ + $H_2$, where the TST and "accurate" data are taken from Allison and Truhlar. The comparisons with ECK and GP are given in Figure S5. The level of agreement with the accurate results is slighter lower in the tunneling region than that seen above for O + $CH_4$. However, we note that "accurate" in this case is ICVT/LAT. Quantum rate constants are given by Allison and Truhlar at 300 and 400 K. At the former temperature, the ratio of ICVT/LAT to the quantum result is 1.3, and at the latter, this ratio is 1.11. Therefore, it is likely that below 300 K ICVT/LAT is larger than the accurate QM result and therefore GP would be even more accurate than what would be concluded based on the comparison with ICVT/LAT below 300 K.

In summary, the GP correction to ECK rate constants, which are already significant improvements over TST results, especially in the tunneling region, is successful. This success was demonstrated mostly for collinear reactions (shown in Figure 1). This is the class of reaction where 85% of the GP training was done, and it is worth recalling that the training data set is roughly 1/3 of the size of the test data set. GP predictions of rate constants were made for four reactions in 3D, with very good outcomes. In fact, the GP predictions are semiquantitative for $O(^3P)$ + $CH_4$ and slightly less accurate for O + $H_2$ and $H_2$ + OH.

At first glance, training on collinear reactions and predicting for 3D ones may seem inconsistent as the dynamics in 3D and in collinear 1D are not identical. However, there is precedent in the literature for doing this. If we consider the product of $\kappa_{ECK}(T)\chi(T)$ to be the overall transmission coefficient correcting TST, then we recall that using the exact transmission coefficient for a collinear reaction to correct TST in 3D was reported in the literature in the 1980s.[11,12] This approach was termed "CEQ/TST",[12] and it required knowing the exact rate constant for the collinear reaction. Here we are using ML to obtain this exact collinear transmission coefficient. The "CEQ/TST" correction to 3D TST was tested for the H + $H_2$[11] and $O(^3P)$ + $H_2$[12] reactions. It was shown to be more accurate for the latter reaction than the former one, especially at low temperatures. We believe that this mirrors the results here for H + $CH_4$ and O + $CH_4$ reactions, and the explanation is likely the same. Namely, in 3D, additional, vibrationally adiabatic contributions (dominated by the ground-state adiabatic energy) to the reaction barrier may occur. An example of this is the restricted bending motion of the atom−molecule at the saddle point. Asymptotically, there is free relative rotational motion of the atom and molecule. The size of this additional barrier depends on the mass of the atom and the bending force constant at the TS. Tunneling through this adiabatic barrier is thus possible and can lead to an overall increase in the rate constant. One successful way to deal with tunneling through these adiabatic barriers is to add the adiabatic energy to the bare potential and to perform an exact quantum calculation of $k(T)$ on the effective potential. We termed this the "CEQB" approach.[11,12] It also goes under the more general rubric of "reduced dimensionality" quantum approaches.[13−15,46,47] For the present applications, we are using GP to effectively do CEQ/TST-quality calculations of $k(T)$. That the approach works better for O + $CH_4$ than H + $CH_4$ can now be understood based on the differences in ground-state adiabatic barriers, and the ground-state adiabatic addition to the potential barrier is roughly 1.5 kcal/mol for H + $CH_4$ compared to roughly half that for the O + $CH_4$ reaction. For O + $H_2$ and OH + $H_2$, where the accuracy of the GP results falls between H + $CH_4$ and O + $CH_4$, inclusion of the adiabatic barrier would probably improve the accuracy. However, we do not propose doing this in future work, which we briefly describe below in the concluding remarks.

Finally, we remark on outliers. These are predictions from ML that disagree substantially with known data used in testing. This is a common but hopefully rare occurrence that can generally be traced to a lack of sufficient data in the training in the vicinity of the prediction input. The F + $H_2$ and F + $D_2$ were labeled as outliers in the discussion of RMS errors shown in Figure 1. The outlier property was that the RMS for the GP prediction is somewhat larger than the ECK one, which is nearly identical to the TST RMS. These reactions are highly exoergic and with a small barrier (about 1.5 kcal/mol on the PES used). There is no training data with these characteristics, and therefore, these outliers are not surprising. We also described the GP prediction for Cl + DH as an outlier. Specifically, the TST rate constant is more accurate than ECK and GP, which are nearly equal. The explanation for this is not obvious, and in fact, this anomalously accurate result for TST was noted by Allison and Truhlar,[8] who used the term "fortuitous" for this case. In any case, these outliers are still in reasonable agreement with accurate results.

To conclude, the present work is the first effort (to the best of our knowledge) to apply ML to represent thermal rate constants of bimolecular chemical reactions. Because ML requires a large data set to train on, we used exact quantum rate constant calculations, mostly for collinear atom + diatom, A + BC, reactions. The training was done on the correction to the product of the conventional TST times the Eckart transmission coefficient to reproduce the exact rate constant. Training was done using GP regression on 13 reactions (defined by a variety of A + BC systems and PESs) spanning a range of reaction barrier heights and reaction endo/exoergicities. The descriptors of $\chi$ ranged from the three parameters of the Eckart transmission coefficient to these plus the skew angle to finally including the vibrational frequency of the BC molecule. Training was done for 13 reactions, and testing was carried out for 40 reactions. The success of the training was demonstrated in terms of the average RMS difference between the exact and trained rate constant. In more than 81% of the total examples and 75% of the test examples,



DOI: 10.1021/acs.jpclett.9b01810
J. Phys. Chem. Lett. 2019, 10, 5250−5258

the GP result was more accurate than either TST or ECK, and on average, it was within a factor of 1.6 of the accurate answer for the full set of reactions. Specific examples were considered in detail, namely, Cl + HCl, trained in 1D and tested for 3D, and the 3D reactions O + $CH_4$ → OH + $CH_3$ and $H_2$ + OH → H + $H_2O$ and O + $CH_4$, and compared to accurate quantum rate constants.

We investigated several sets of descriptors. Because the training is done on corrections to Eckart tunneling, we chose as a minimal set of descriptors the three parameters used in Eckart tunneling. Overall results based on these (cf. Table 2) are good, but improvements were found by adding two additional descriptors, the skew angle and the symmetric stretch vibrational frequency of the reactant. The effect of the skew angle is not surprising as this angle characterizes mass combination, which for heavy−light−heavy systems can lead to important dynamical effects such as corner-cutting and recrossing. A further improvement was found by adding a symmetric stretching frequency, which in an adiabatic theory correlates with a frequency of a TS normal mode. Future work using the ML approach described here can be extended in a number of ways. For example, the current database of exact corrections to ECK, given in the SI, could be partitioned in other ways for training and testing. Similarly, the entire database (see Table S3) could be used for training and then ready for predictions on new reactions. Also, the current database is not comprehensive and could be extended by training on polyatomic reactions, provided accurate rate constant data exists. Ideally, future training will be done on a large and diverse group of 3D reactions for which accurate, ideally exact rate constants are available. This could be from a combination of theory or experiment. For high temperatures, quasiclassical trajectory calculations of rate constants can provide accurate results, and therefore, these could be used in such an extended data set.

Finally, we note recent work using ML, specifically Neural Networks, to model state-to-state reactions cross sections, obtained using quasiclassical trajectory calculations. The goal is to use ML training to permit the prediction of a set of cross sections that can be used to model high-temperature kinetics. The approach was successfully applied to the N + NO → O + $N_2$ reaction.[48]

## COMPUTATIONAL DETAILS

From Table 1 in Allison and Truhlar,[8] it is possible, using standard heats of formation,[39] to calculate the parameters $\alpha_1$, $\alpha_2$, and $u^*$ needed to perform a correction to TST based on the Eckart model.[1−3] In order to evaluate this correction, we used the approach by Johnston and Heicklen[3] as modified by Brown.[28] Integration of the transmission coefficient over energy was performed using the Mathematica **NIntegrate** function.[49] In addition to the three parameters just mentioned, we also tabulated $\beta$, the skew angle for the reaction, as well as $\omega_e$(B−C), the harmonic frequency of the B−C bond in the reaction A + BC → AB + C, available from the NIST WebBook.[39] A summary of the parameters used in this study can be found in Table S1.

GP regression was performed using the Mathematica **Predict** routine,[49] with the following options: "EstimationMethod" → "MaximumPosterior", "CovarianceType" → "SquaredExponential", AssumeDeterministic → True. Nearly identical results were obtained with "OptimizationMethod" set to "FindMinimum" or "SimulatedAnnealing".

Training data is based on rate constants reported in Allison and Truhlar.[8] The training data set for the five-parameter version is provided in Table S2. The training time for the five-parameter version was 1.7 s in Mathematica. The training set was 13 reaction/potential surface combinations with results at 8 temperatures in each, or 104 examples. Thus, the training, which involves the inversion of a 104 × 104 matrix, is quite fast.

GP regression was also done using the open source Python code[50] with the same squared exponential covariance matrix as above. The training results shown in Table S2 are to within four significant figures the same.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.9b01810.

Parameters for the reactions in Table 1, the complete set combinations of training set parameters, along with the target values of $\chi$, and the predicted results, the complete set combinations of test set parameters, along with the target values of $\chi$, and the predicted results, and comparisons between accurate calculations, TST, ECK, and GP results for the various reactions (PDF)

## AUTHOR INFORMATION


### Corresponding Authors

*E-mail: plh2@cornell.edu.
*E-mail: apurba.nandi@emory.edu.
*E-mail: jmbowma@emory.edu.

### ORCID

Paul L. Houston: 0000-0003-2566-9539
Apurba Nandi: 0000-0002-6191-5584
Joel M. Bowman: 0000-0001-9692-2672


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Johannes Kästner for sending the corrected TST rate constants for the OH + $H_2$ reaction and Dong Hui Zhang for sending the accurate coupled channel ones. J.M.B. thanks the National Science Foundation (CHE1463552) for funding.

## REFERENCES


(1) Eckart, C. The Penetration of a Potential Barrier by Electrons. *Phys. Rev.* **1930**, *35*, 1303−1309.
(2) Johnston, H. S.; Rapp, D. Large Tunnelling Corrections in Chemical Reaction Rates.1 II. *J. Am. Chem. Soc.* **1961**, *83*, 1−9.
(3) Johnston, H. S.; Heicklen, J. Tunneling Corrections for Unsymmetrical Eckart Potential Energy Barriers. *J. Phys. Chem.* **1962**, *66*, 532−533.
(4) Barker, J. R.; Nguyen, T. L.; Stanton, J. F. Kinetic Isotope Effects for Cl + $CH_4$ → HCl + $CH_3$ Calculated Using Ab Initio Semiclassical Transition State Theory. *J. Phys. Chem. A* **2012**, *116*, 6408−19.
(5) Nguyen, T. L.; Stanton, J. F.; Barker, J. R. Ab initio Reaction Rate Constants Computed using Semiclassical Transition-State Theory: HO + $H_2$ → $H_2O$ + H and Isotopologues. *J. Phys. Chem. A* **2011**, *115*, 5118−26.
(6) da Silva, G.; Cole, J. A.; Bozzelli, J. W. Thermal Decomposition of the Benzyl Radical to Fulvenallene ($C_7H_6$ + H). *J. Phys. Chem. A* **2009**, *113*, 6111−20.




DOI: 10.1021/acs.jpclett.9b01810
*J. Phys. Chem. Lett.* 2019, 10, 5250−5258

(7) Cavallotti, C.; Pelucchi, M.; Georgievskii, Y.; Klippenstein, S. J. EStokTP: Electronic Structure to Temperature- and Pressure-Dependent Rate Constants-A Code for Automatically Predicting the Thermal Kinetics of Reactions. *J. Chem. Theory Comput.* **2019**, *15*, 1122−1145.
(8) Allison, T. C.; Truhlar, D. G. In *Modern Methods for Multidimensional Dynamics Computations in Chemistry*; Thompson, D. L., Ed.; World Scientific: Singapore, 1998; pp 618−712.
(9) Schatz, G. C.; Kuppermann, A. Quantum Mechanical Reactive Scattering for Three-Dimensional Atom Plus Diatom Systems. II. Accurate Cross Sections for H+$H_2$. *J. Chem. Phys.* **1976**, *65*, 4668−4692.
(10) Marcus, R. A.; Coltrin, M. E. A New Tunneling Path for Reactions such as H+$H_2$ →$H_2$+H. *J. Chem. Phys.* **1977**, *67*, 2609−2613.
(11) Bowman, J. M. *Advances in Chemical Physics*; John Wiley and Sons, Ltd., 1985; pp 115−167.
(12) Bowman, J. M.; Wagner, A. F. In *The Theory of Chemical Reaction Dynamics*; Clary, D., Ed.; NATO ASI Series (Series C: Mathematical and Physical Sciences); Springer: Dordrecht, The Netherlands, 1986; Vol. *170*; Chapter 6, pp 129−164.
(13) Bowman, J. M. Reduced Dimensionality Theory of Quantum Reactive Scattering. *J. Phys. Chem.* **1991**, *95*, 4960−4968.
(14) Althorpe, S. C.; Clary, D. C. Quantum Scattering Calculations on Chemical Reactions. *Annu. Rev. Phys. Chem.* **2003**, *54*, 493−529.
(15) von Horsten, H. F.; Banks, S. T.; Clary, D. C. An Efficient Route to Thermal Rate Constants in Reduced Dimensional Quantum Scattering Simulations: Applications to the Abstraction of Hydrogen from Alkanes. *J. Chem. Phys.* **2011**, *135*, 094311.
(16) Miller, W. H. Semiclassical Limit of Quantum Mechanical Transition State Theory for Nonseparable Systems. *J. Chem. Phys.* **1975**, *62*, 1899−1906.
(17) Miller, W. H.; Hernandez, R.; Handy, N. C.; Jayatilaka, D.; Willetts, A. Ab initio Calculation of anharmonic Constants for a Transition State, with Application to Semiclassical Transition State Tunneling Probabilities. *Chem. Phys. Lett.* **1990**, *172*, 62−68.
(18) Hernandez, R.; Miller, W. H. Semiclassical Transition State Theory. A New Perspective. *Chem. Phys. Lett.* **1993**, *214*, 129−136.
(19) Wagner, A. F. Improved Multidimensional Semiclassical Tunneling Theory. *J. Phys. Chem. A* **2013**, *117*, 13089−100.
(20) Clary, D. C. Spiers Memorial Lecture Introductory lecture: quantum dynamics of chemical reactions. *Faraday Discuss.* **2018**, *212*, 9−32.
(21) Craig, I. R.; Manolopoulos, D. E. Chemical Reaction Rates from Ring Polymer Molecular Dynamics. *J. Chem. Phys.* **2005**, *122*, 084106.
(22) Suleimanov, Y. V.; de Tudela, R. P.; Jambrina, P. G.; Castillo, J. F.; Saez-Rabanos, V.; Manolopoulos, D. E.; Aoiz, F. J. A Ring Polymer Molecular Dynamics Study of the Isotopologues of the H+$H_2$ Reaction. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3655−3665.
(23) Richardson, J. O. Perspective: Ring-Polymer Instanton Theory. *J. Chem. Phys.* **2018**, *148*, 200901.
(24) Miller, W. H.; Schwartz, S. D.; Tromp, J. W. Quantum Mechanical Rate Constants for Bimolecular Reactions. *J. Chem. Phys.* **1983**, *79*, 4889−4898.
(25) Wu, T.; Werner, H.-J.; Manthe, U. First-Principles Theory for the H + $CH_4$ → $H_2$ + $CH_3$ Reaction. *Science* **2004**, *306*, 2227−2229.
(26) Welsch, R.; Manthe, U. Reaction Dynamics with the Multi-Layer Multi-Configurational Time-Dependent Hartree Approach: H + $CH_4$ →$H_2$ + $CH_3$ Rate Constants for Different Potentials. *J. Chem. Phys.* **2012**, *137*, 244106.
(27) Suleimanov, Y. V.; Collepardo-Guevara, R.; Manolopoulos, D. E. Bimolecular Reaction Rates from Ring Polymer Molecular Dynamics: Application to H + $CH_4$ → $H_2$ + $CH_3$. *J. Chem. Phys.* **2011**, *134*, 044131.
(28) Brown, R. L. A Method of Calculating Tunneling Corrections for Eckart Potential Barriers. *J. Res. Natl. Bur. Stand.* **1981**, *86*, 357−359.
(29) Wolfsberg, M.; et al. *Isotope Effects: In the Chemical, Geological, and Bio Sciences*; Spinger: Dordrecht, Heidelberg, New York, 2010; pp 181−201.
(30) Espinosa-Garcia, J.; Garcia-Bernaldez, J. C. Analytical Potential Energy Surface for the $CH_4$+ $O(^3P)$→ $CH_3$ + OH Reaction. Thermal Rate Constants and Kinetic Isotope Effects. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2345−2351.
(31) Li, Y.; Suleimanov, Y. V.; Yang, M.; Green, W. H.; Guo, H. Ring Polymer Molecular Dynamics Calculations of Thermal Rate Constants for the $O(^3P)$ + $CH_4$ → OH + $CH_3$ Reaction: Contributions of Quantum Effects. *J. Phys. Chem. Lett.* **2013**, *4*, 48−52.
(32) Li, J.; Chen, J.; Zhao, Z.; Xie, D.; Zhang, D. H.; Guo, H. A Permutationally Invariant Full-Dimensional Ab Initio Potential Energy Surface for the Abstraction and Exchange Channels of the H + $CH_4$ System. *J. Chem. Phys.* **2015**, *142*, 204302.
(33) Meng, Q.; Chen, J.; Zhang, D. H. Communication: Rate Coefficients of the H + $CH_4$ → $H_2$ + $CH_3$ Reaction from Ring Polymer Molecular Dynamics on a Highly Accurate Potential Energy Surface. *J. Chem. Phys.* **2015**, *143*, 101102.
(34) Chen, J.; Xu, X.; Xu, X.; Zhang, D. H. A Global Potential Energy Surface for the $H_2$ + OH → $H_2$ + H Reaction using Neural Networks. *J. Chem. Phys.* **2013**, *138*, 154301.
(35) Meisner, J.; Kastner, J. Reaction Rates and Kinetic Isotope Effects of $H_2$ + OH → $H_2O$ + H. *J. Chem. Phys.* **2016**, *144*, 174303.
(36) Sun, P.; Zhang, Z.; Chen, J.; Liu, S.; Zhang, D. H. Well Converged Quantum Rate Constants for the $H_2$ + OH → $H_2O$ + H Reaction via Transition State Wave Packet. *J. Chem. Phys.* **2018**, *149*, 064303.
(37) Welsch, R. Low-Temperature Thermal Rate Constants for the Water Formation Reaction $H_2$ + OH from Rigorous Quantum Dynamics Calculations. *Angew. Chem.* **2018**, *130*, 13334−13337.
(38) Welsch, R. Rigorous Close-Coupling Quantum Dynamics Calculation of Thermal Rate Constants for the Water Formation Reaction of $H_2$ + OH on a High-Level PES. *J. Chem. Phys.* **2018**, *148*, 204304.
(39) Lindstrom, P. J. *NIST Standard Reference Database Number 69*. https://webbook.nist.gov/chemistry/#Documentation (2018).
(40) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. *Int. J. Quantum Chem.* **2015**, *115*, 1051−1057.
(41) 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.
(42) Qu, C.; Yu, Q.; Van Hoozen, B. L.; Bowman, J. M.; Vargas-Hernández, R. A. Assessing Gaussian Process Regression and Permutationally Invariant Polynomial Approaches To Represent High-Dimensional Potential Energy Surfaces. *J. Chem. Theory Comput.* **2018**, *14*, 3381−3396.
(43) Rasmussen, C. E.; Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*; The MIT Press, 2005.
(44) Schatz, G. C.; Amaee, B.; Connor, J. N. L. A Centrifugal-Sudden Distorted-Wave Study of Isotope Effects for the Reactions Cl + HCl→ClH+Cl and Cl+DCl→ClD+Cl. *J. Chem. Phys.* **1990**, *93*, 5544−5551.
(45) Meisner, J.; Kastner, J. Erratum: "Reaction Rates and Kinetic Isotope Effects of H2 + OH → H2O + H" [J. Chem. Phys. 144, 174303 (2016)]. *J. Chem. Phys.* **2017**, *147*, 049903.
(46) Wang, D.; Bowman, J. M. A Reduced Dimensionality, Six-Degree-of-Freedom, Quantum Calculation of the H+$CH_4$ → $H_2$+$CH_3$ Reaction. *J. Chem. Phys.* **2001**, *115*, 2055−2061.
(47) Banks, S. T.; Clary, D. C. Reduced Dimensionality Quantum Dynamics of Cl + $CH_4$→HCl + $CH_3$ on an Ab Initio potential. *Phys. Chem. Chem. Phys.* **2007**, *9*, 933−43.
(48) Koner, D.; Unke, O. T.; Boe, K.; Bemish, R. J.; Meuwly, M. Exhaustive State-to-State Cross Sections for Reactive Molecular Collisions from Importance Sampling Simulation and a Neural Network Representation. *J. Chem. Phys.* **2019**, *150*, 211101.


DOI: 10.1021/acs.jpclett.9b01810
*J. Phys. Chem. Lett.* 2019, 10, 5250−5258

(49) *Mathematica*, version 12.0; Wolfram Research, Inc.: Champaign, IL, 2019.
(50) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.



DOI: 10.1021/acs.jpclett.9b01810
*J. Phys. Chem. Lett.* 2019, 10, 5250−5258