PAPER

Check for updates

Potential energy surfaces: Δ-machine learning from analytical functional forms

Cipriano Rangel [iD] *[a] and Joaquin Espinosa-Garcia [iD] [b]

Delta-machine learning (Δ-ML) is a highly cost-effective approach for developing high-level potential energy surfaces (PESs) from a large number of low-level configurations. In particular, the high flexibility of the analytical potential energy surface developed previously by our group is exploited to efficiently sample points from the low-level data set and, using information from the highly accurate permutation invariant polynomial neural network (PIP-NN) surface, construct the Δ-ML PES. This approach is applied to the well-known H + CH$_4$ hydrogen abstraction reaction. In order to test the validity and accuracy of the approach to describe this polyatomic system, kinetic studies using the variational transition state with multidimensional tunneling corrections and dynamic studies on the deuterated reaction, H + CD$_4$, using quasiclassical trajectory calculations were performed on three surfaces. The delta-machine learning approach reproduces the kinetics and dynamics information of the high-level surface, showing its efficiency in describing multidimensional polyatomic systems.

Introduction

The elucidation and characterization of potential energy surfaces (PESs) are fundamental in computational chemistry and physics, as they provide the foundation for understanding molecular interactions, reactivity, and dynamics in polyatomic systems with multiple degrees of freedom. Accurate description requires high-level *ab initio* calculations with electron correlation, which are extremely demanding computationally, especially for large systems. Machine learning (ML) methods offer a means to significantly reduce computational time.[1] In recent years, the application of ML in the development of PESs has gained considerable attention,[2–27] based on different strategies, such as neural networks (NNs),[2–17] Gaussian process regression (GPR),[18–22] or permutationally invariant polynomials (PIPs).[23,24] Recent comprehensive reviews provide a broad overview of machine learning methods applied to atomistic simulations, highlighting strategies for efficient PES generation that balance computational cost and accuracy, and situate delta-machine learning methods within this evolving landscape. Zhu *et al.*[28] applied Δ-machine learning to construct a PES, using DFTB as the low-level reference instead of an analytical potential as in the present work, thereby illustrating both methodological similarities and differences relevant to this field. Combined approaches, such as PIP-NNs, have also been successful for polyatomic systems.[25–27] These strategies have revolutionized

the development and refinement of PESs, offering unprecedented accuracy. While artificial NNs have the ability to capture non-linear relationships and intricate patterns, providing a versatile framework for modelling atomic interactions and energy contributions, GPR facilitates uncertainty quantification, enabling researchers to assess the reliability and robustness of predicted energy values across molecular configurations. These strategies have a common denominator: they need a large number of high-level (HL) *ab initio* calculations, which, for polyatomic systems with many degrees of freedom, represents a computational bottleneck.

Complementing the previous methods, the delta-machine learning, Δ-ML, technique[29–32] provides a holistic approach to PES development, emphasizing the integration of dynamical theories with empirical data analysis. The main advantage of this technique is its ability to circumvent expensive HL *ab initio* calculations. In the Δ-ML method, a huge set of points obtained using low-level, LL, electronic structure calculations describe the polyatomic system, which are corrected with a few HL ones,[33,34] using a correction term, ($\Delta V_{\text{HL–LL}}$). Thus, the corrected PES can be written as follows:[33]

$$V_{\text{HL}}^i = V_{\text{LL}}^i + \Delta V_{\text{HL–LL}}^i$$

where the superscript i refers to the ith geometric configuration with energy E. Since the correction term, $\Delta V_{\text{HL–LL}}$, is a slowly varying function of the atomic coordinates, it can be machine learned from a small number of data points, reasonably chosen. Thus, it represents an efficient scheme to reduce the computational cost for PES construction. The main idea consists of developing a high accuracy PES as the sum of a low

[a] *Area de Química Orgánica, Spain. E-mail: ciprira@unex.es*
[b] *Area de Quimica Fisica and Instituto de Computacion Cientifica Avanzada, Universidad de Extremadura, 06071 Badajoz, Spain*

accuracy surface based on a huge set of points and a correction term equivalent to the difference for high and low accuracy surfaces based on a reduced and judiciously chosen set of points. One should note that, to some extent, this scheme is reminiscent of previous methods such as interpolated correction,[35] IRC-Max,[36] or even the ONIOM method,[37] and the first two focus on the minimum energy path. With ML, researchers can now achieve the development of PESs for polyatomic systems with chemical accuracy (*ca.* 1 kcal mol^{-1}), which was unthinkable just a decade ago. Typically, in the delta-ML technique, different LL methods have been used, such as Hartree–Fock (HF), complete active space self-consistent field (CASSCF), or density functional theory (DFT). However, in the present paper, we propose a different alternative, using LL as the analytical PESs previously built by our group, which did not reach the chemical accuracy in many cases. This represents a great advantage since these analytical surfaces were developed as a full-dimensional problem, representing the entire reactive system, from reactants to products, with an asymptotic behaviour in the entrance and exit channels. The efficacy and high-fidelity of this approach will be demonstrated using the well-known H + CH$_4$ hydrogen abstraction reaction because much kinetics and dynamics information is known theoretically and experimentally. Theoretically, this gas-phase reaction has a long history in our research group, beginning in 1996 and finishing in 2009, and so the PES-1996,[38] PES-2002,[39] and PES-2008[40] surfaces were developed based on valence-bond molecular mechanics, VB-MM-type, analytical functional forms, correcting in each improvement the limitations found in the previous developments. Bowman and co-workers developed a series of potential energy surfaces using the PIP strategy,[23,24,41] fitting about 21 000 electronic energies calculated using the spin-restricted coupled-cluster method RCCSDT with a large basis aug-cc-pVTZ, in brief, RCCSD(T)/aug-cc-pVTZ, which are known as ZBBx surfaces. Later, Zhang and co-workers[42] developed the ZFWCZ surface in 2011, using the modified Shepard interpolation scheme and energies calculated at the UCCSD(T)/aug-cc-pVTZ level. The most recent and accurate surface was constructed in 2015 by Li and co-workers,[43] which was a PIP-NN surface fitted to about 63 000 *ab initio* points at the UCCSD(T)-F12a/AVT level. In the present paper, we use V_{LL} as the PES-2008 surface,[40] which is an analytical surface fitted to high-level *ab initio* calculations, and V_{HL} energies were obtained from the PIP-NN surface developed by Li *et al.*[43] This Δ-ML surface was compared with the original PIP-NN surface[43] and the PES-2008 surface.[40]

Theoretical tools

(A) Delta-machine learning process

As previously noted, HL energies were obtained from the PIP-NN PES constructed by Li *et al.*,[43] which was based on ~63 000 points calculated at a high *ab initio* level, UCCSD(T)-F12a/AVTZ, using the PIP-NN strategy. This surface presents a small fitting error, ~5.1 meV, 0.12 kcal mol^{-1}, or 42 cm^{-1}, and so it was

Table 1 Relative energies with respect to reactants for the H + CH$_4$ → H$_2$ + CH$_3$ hydrogen abstraction reaction. All energy values are in kcal mol^{-1}

Method	ΔE_R	$\Delta E^{\neq}$	Ref.
CCSD(T)/cc-pVTZ	2.9	15.1	40
CCSD(T)/aug-cc-pVTZ	2.8	14.8	41
CCSD(T)/aug-cc-pVQZ	2.9	14.9	44
Best estimate		14.8	45
CCSD(T)-F12a/aug-cc-pVTZ	2.7	14.6	43
CCSD(T)-F12a/aug-cc-pVTZ	2.9	14.8	Present work
+Core	3.0	14.9	Present work
+Core + relativistic	2.8	14.7	Present work
PES-2008	2.9	15.0	40
PIP-NN	2.8	14.7	43

used to test the accuracy of the Δ-ML PES. The PIP-NN surface is possibly one of the most accurate representations of the gas phase H + CH$_4$ hydrogen abstraction reaction; however, these authors recognised that errors of 90–180 cm^{-1} can be expected in the relative energies because some small corrections, such as the core correlation and the relativistic effect, were not taken into account in the UCCSD(T)-F12a/AVTZ *ab initio* level. To estimate the influence of these corrections, we have performed new calculations here by analysing the energy of the reaction and the barrier height, and the results are shown in Table 1, together with other theoretical estimations for comparison. The estimations of the barrier height are in the narrow range of 14.6–15.1 kcal mol^{-1}, *i.e.*, small differences of only 0.5 kcal mol^{-1}, with the most recent and accurate values in the narrower range of 14.6–14.9 kcal mol^{-1}. We have estimated that the effects of the core correlation and the relativistic effects on the barrier height are only 0.1 kcal mol^{-1} and ~35 cm^{-1}. With these corrections, the barrier height is 14.7 kcal mol^{-1}, in agreement with Pu and Truhlar's best estimate[45] and the PIP-NN value,[43] improving the results obtained with PES-2008.[40]

The development of a high-level (HL), full-dimensional potential energy surface (PES) is challenging and computationally demanding. To improve the efficiency of the PES description, a correction approach can be employed, where a low-level (LL) energy calculation is adjusted by the difference between the LL and HL calculations (ΔHL–LL), as shown below:

$$V_{HL} = V_{LL} + \Delta V_{HL-LL}$$

In this work, a new strategy, termed Δ-VB-MM-PES, was developed. This strategy employs an analytical, full-dimensional potential energy surface based on valence bond-molecular mechanics (VB-MM) developed previously in our group[40] as the LL energy calculation, PES-2008. For the HL calculations, the permutation invariant polynomial neural network (PIPNN) developed by Li *et al.*[43] was selected, which was fitted using the UCCSD(T)-F12a/AVTZ level of theory. Thereafter, an artificial neural network (ANN) model was used to fit the ΔV_{HL-LL} to construct the HL PES (Δ-ML).

(a) Energy calculations. In order to obtain a comprehensive description of the potential energy surface (PES), a set of 100 000 geometries were sampled, covering a wide variety of molecular configurations along the PES with a broad configuration

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19204–19215 | **19205**

space, and C–Ha and Ha–Hb distances reaching a maximum of 10 Å, ensuring the separation of the reactant and product in the asymptotic regions. The energy was calculated at every geometry at a low level (LL-PES and PES-2008). Here, the distances C–Ha and Ha–Hb represent, respectively, the bond length between the central carbon atom and the hydrogen atom (Ha) directly bonded to it that is abstracted, and the distance between this hydrogen atom (Ha) and the attacking atom (Hb) involved in the abstraction reaction. From the set of geometries calculated at the low level (LL), 30 000 geometries were selected, with the focus being on the reactants, products, saddle points, and reaction swath regions according to a multistep sampling strategy designed to ensure thorough coverage of the most relevant regions of the PES.

Specifically, initial broad scans of the configuration space were performed, followed by full optimisations of the reactant, product, and saddle point structures. Using the optimised saddle point as a reference, an intrinsic reaction coordinate (IRC) calculation was carried out to map the minimum energy path (MEP) connecting the reactants and products. Additional scans were then performed along this IRC path, as well as around the key stationary points (reactants, products, and transition state), densely sampling the regions critical for the reaction mechanism. This approach guaranteed increased density of configurations in the reactant and product wells, the saddle point region, and along the reaction swath, while maintaining a representative spread of configurations in the global PES.

Fig. 1 depicts the distances C–Ha and Ha–Hb at each point, with the maximum reaching 6 Å. At each geometry, a single point high-level (HL) calculation was performed with the PIP-NN-PES.

(b) Fitting model. An artificial neural network (ANN) model based on a multilayer perceptron (MLP) implemented in scikit-learn[46] was used in order to fit the difference between the HL and the LL of energies measured over the 30 000 data points. The feature set used in the neural network fitting consists of the Cartesian coordinates of all atoms in each geometry. Hyperparameters (number of neurons, layers, *etc.*) were chosen manually. This model was structured with two hidden layers, with the first hidden layer consisting of 100 neurons and the second, larger layer consisting of 1000 neurons. The loss function, which is minimised to obtain the ANN model, represents the difference between the predicted energy from the neural network model and the true energy difference between the LL and HL calculations. A maximum of 1000 iterations was set, meaning that the network would update its weights up to 1000 times based on the error at each step. A convergence tolerance of 1×10^{-4} was used to define the stopping criterion.

The efficacy of the model can be evaluated by analysing the root mean square error (RMSE) values, which decrease from 0.5447 kcal mol^{-1} when the difference between V_{HL} and V_{LL} energies is found to be 0.0263 kcal mol^{-1} after implementing the correction with the model V_{HL}-Δ-ML over the grid employed to fit the model.

The difference between PES-2008 (LL) and PIP-NN (HL) before and after applying the ANN model to obtain the Δ-ML energy is shown in Fig. 2.

A further significant evaluation of the fitting model can be conducted by examining the geometries obtained from the minimum energy path (MEP). These geometries are of particular importance in the kinetics of the reaction. As illustrated in Fig. 3, the PES-2008 energies are plotted against the geometries obtained from the MEP, along with single points at the PIP-NN and the PES-2008 + Δ-ML energies. In this instance, the root mean square error (RMSE) changes from 0.6405 kcal mol^{-1} to 0.1200 kcal mol^{-1} when the model is applied to the original values of the PES-2008. Note that in Fig. 3, we are drawing the
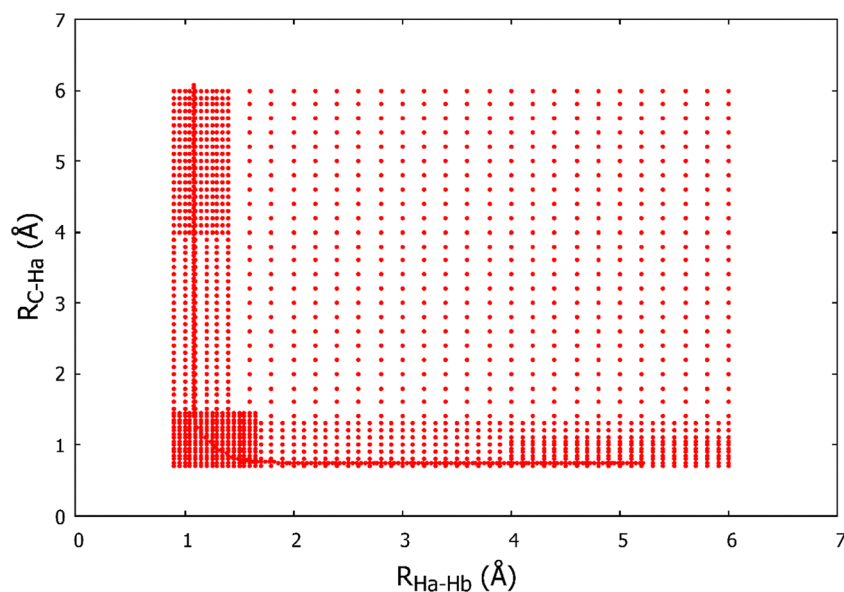
Fig. 1 Scatter plot of the distances R_{C-Ha} and R_{Ha-Hb} in Å for the grid of molecular geometries used to fit the machine learning model. R_{C-Ha} represents the bond length between the central carbon atom and a hydrogen atom Ha in methane, and R_{Ha-Hb} corresponds to the distance between the abstracted hydrogen Ha and the incoming hydrogen atom Hb.

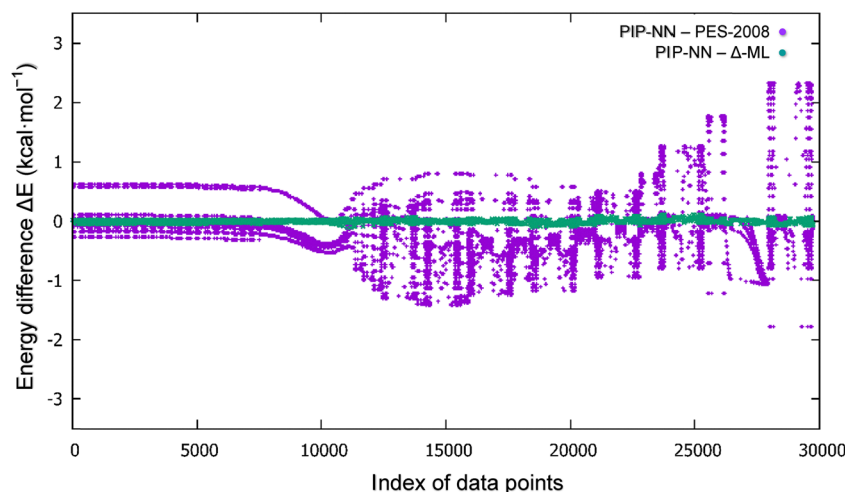**19206** | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19204–19215

This journal is © the Owner Societies 2025

Fig. 2 Energy differences (in kcal mol^{-1}) between the PIP-NN surface and the PES-2008 surface (purple dots) and between the PIP-NN surface and the Δ-ML corrected surface (green dots), evaluated over the grid of data points used for fitting.

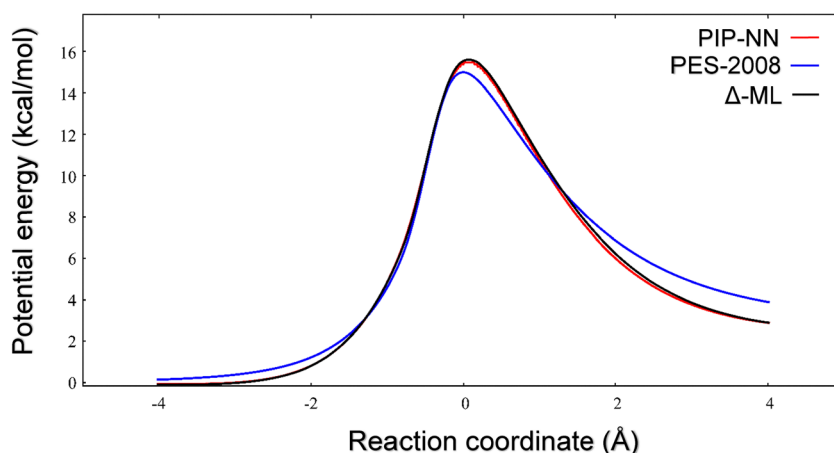**Fig. 3** Minimum energy paths for the PES-2008 (blue line), PIP-NN (red line), and Δ-ML (black line). Note that the PIP-NN and PES-2008 + Δ curves were obtained as single-point calculations at these levels, based on optimized geometries at the PES-2008 level.

MEP optimized at the PES-2008, and on these geometries, single-point calculations at the PIP-NN and PES-2008 + Δ-ML levels were performed.

To evaluate the efficacy of the procedure, a novel grid was selected in the reaction swath, with a limit of 60.0 kcal mol^{-1} in terms of relative energy (*i.e.*, with respect to the reactant limit). A total of 10 000 points were evaluated at the PES-2008, PIP-NN, and PES-2008 + Δ-ML levels. The contour plot obtained with the PES-2008 (upper panel), PIP-NN (middle panel), and PES-2008 + Δ-ML (bottom panel) is depicted in Fig. 4. When the Δ-ML correction is applied, the description of the saddle point region is improved, and therefore, a better description of the rate constant is expected. Furthermore, the root mean square error (RMSE) decreased from 0.6913 to 0.0673 kcal mol^{-1}.

(B) Kinetics and dynamics computational details

In the present work, for the kinetics study, to explore the temperature dependence of the rate constant, the VTST/MT, *i.e.*, variational transition state theory, was employed based on the PES-2008, PIP-NN, and Δ-ML surfaces, and the results were compared with the available experimental information.[47,48] In the temperature range of 200–2500 K, the canonical version of the VTST (CVT),[49,50] with multidimensional tunneling corrections from the microcanonical optimized multidimensional tunneling approach (μOMT) was used,[51] in brief, CVT/μOMT. In this theory, it was assumed that the vibrational modes are separable harmonic oscillators where the normal modes were computed using a system of redundant internal coordinates.[52] The four hydrogen equivalent paths were considered with a sigma factor of 4. The kinetics calculations were performed using the POLYRATE code,[53] interfaced with three present surfaces.

For the dynamics study, quasi-classical trajectory (QCT) calculations were used on three surfaces at a collision energy of 1.5 eV for the H + CD$_4$ → HD + CD$_3$ reaction because more experimental information is available for this deuterated reaction.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19204–19215 | **19207**
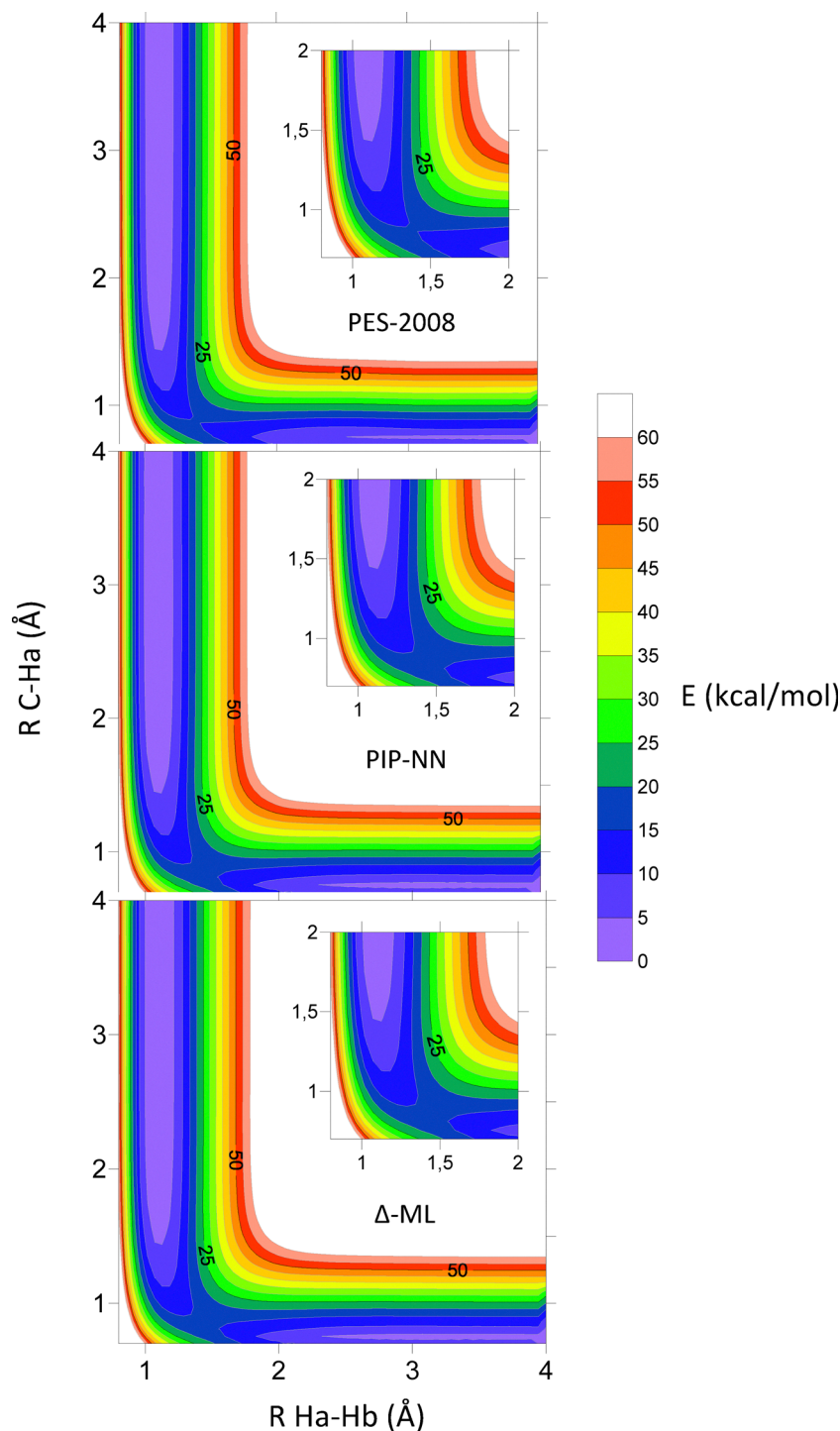
Fig. 4 Equipotential contour plots of the H + CH₄ → H₂ + CH₃ reaction in the saddle point region for the three surfaces: (top) PES-2008, (middle) PIP-NN, and (bottom) Δ-ML. Axes correspond to R_{C-Ha} (Å) and R_{Ha-Hb} (Å). Colors indicate the potential energy in kcal mol⁻¹. The enlargement highlights the region near the transition state. Axis labels and units are identical for all panels.

The maximum impact parameter, b_{max}, was obtained at this energy for each surface, which was obtained by running small batches of trajectories until no reactive trajectories were found. At this energy, the b_{max} value is 2.0 Å. For each surface, 1×10^5 trajectories were run, where the C–H separation at the beginning and the end of the trajectory was fixed at 15.0 Å, with

a propagation step of 0.1 fs. The remaining initial conditions were randomly selected from a Monte Carlo sampling, following the VENUS code.[54,55] Note that due to its classical nature, the QCT calculations present two important limitations. Firstly, quantum effects are not considered, but since the collision energy is high, the tunneling contribution can be considered

19208 | Phys. Chem. Chem. Phys., 2025, 27, 19204–19215

This journal is © the Owner Societies 2025

small or even negligible. Secondly, the zero-point energy (ZPE) violation problem appears, associated with the fact that some trajectories end with vibrational energy below the ZPE of the products. This issue was considered using two alternatives, which remain intrinsically *ad hoc*: (i) considering all reactive trajectories (the All approach), or only reactive trajectories with the vibrational energy of each product above their respective ZPEs (the DZPE or double ZPE approach), HD and CD_3. Note that this latter approach, though physically reasonable, drastically reduces the number of reactive trajectories to be considered in the final analysis.

Results and discussion

Before beginning the analysis of the theoretical results, two points of caution must be considered for a rigorous comparison: (1) when different surfaces are compared using the same theoretical tools, the quality and accuracy of the PES is tested, but (2) when theoretical results are compared with experiments, both the PES and the theoretical tools are simultaneously tested.

(1) Kinetics results

For the per-protio reaction, $H + CH_4 \rightarrow H_2 + CH_3$, Fig. 5 shows the Arrhenius plots of the corresponding forward rate constants in the temperature range of 200–2500 K for three potential energy surfaces, PES-2008, PIP-NN and the present Δ-ML, using the same kinetic tool, canonical variational transition state theory (CVTST), where the tunneling correction at low temperatures was included by using the semiclassical microcanonical optimized multidimensional tunneling approach, μOMT, in brief, CVT/μOMT. Therefore, we are testing the accuracy of

the PES and the validity of the Δ-ML approach, comparing its kinetics results with the PIP-NN and PES-2008 levels. In the temperature range of 300–2500 K, the Δ-ML approach reasonably reproduces the PIP-NN PES, improving the agreement with respect to the PES-2008 approach. However, the improvement is better in the low-temperature range, 200–300 K. Thus, while the PES-2008 approach underestimates the PIP-NN results by a factor of about 4, the Δ-ML approach reproduces the PIP-NN rate constants. When the comparison theory/experiment is considered, we observed good agreement in the high-temperature regime for the three PESs, but at low temperatures, while the PES-2008 surface underestimates the experimental evidence by a factor of about 4; the PIP-NN and the Δ-ML surfaces reproduce the experimental behaviour.

Let us now analyse this behaviour considering separately the high-temperature regime, where the recrossing effects are predominant, and the low-temperature regime, where the tunneling corrections play the most important role. The recrossing effects, also known as the "variational effects" in the VTST approach, measure the number of trajectories returning to reactants, and they are dependent on the choice of the dividing surface. In fact, they measure the shift of the maxima of the free energy curve (located in the reaction coordinate s^*) from the saddle point (by definition, located at $s = 0$) and are obtained as the ratio between CVT(s^*) and TST($s = 0$) forward rate constants. Firstly, the PES-2008 and PIP-NN PESs present very different behaviours. While the PES-2008 approach presents factors of about unity in the 500–2500 K temperature range, indicating that the maxima free energy is located close to the saddle point and, consequently, negligible recrossing effects, the PIP-NN approach presents noticeable contributions in this range, from 0.1 to 0.7, which indicates a shift from the
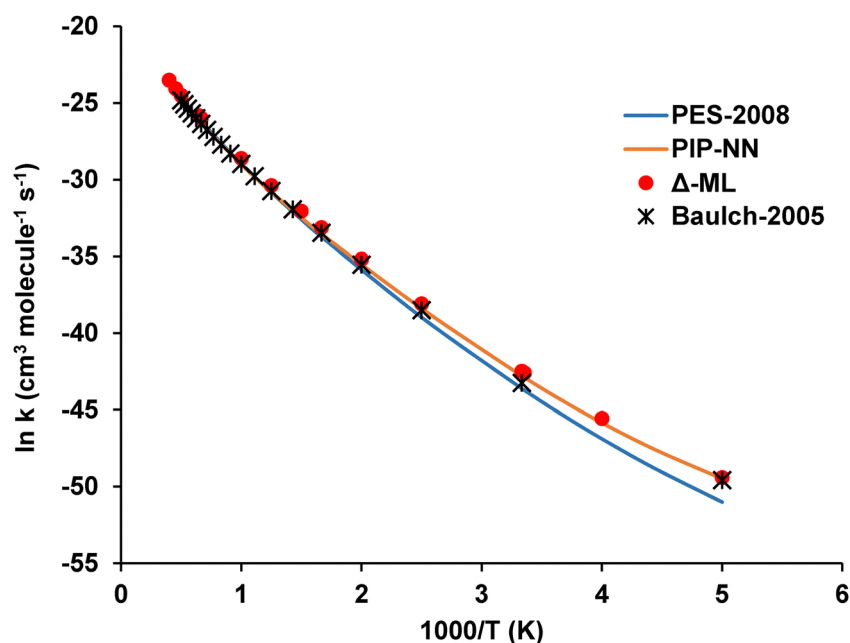

Fig. 5 Overall thermal rate constants of the $H + CH_4$ reaction at temperatures from 200 to 2000 K for the three surfaces, PES-2008, PIP-NN, and Δ-ML, using the CVT/μOMT kinetics model. Experimental measures from the study of Baulch *et al.* 2005 in the range of 300–2000 K, and KIDA2 at 200 K.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19204–19215 | **19209**

saddle point. The Δ-ML surface shows a similar behaviour to the PES-2008 approach, with factors about unity in this temperature range, and, therefore, negligible recrossing. In the opposite extreme, the low temperature regime and tunneling effects, the behaviour of the three PESs is similar to that found for recrossing, with moderate tunneling factors, 6.21×10^2 and 4.32×10^2, for the PES-2008 and Δ-ML surfaces, respectively, at 200 K, as compared to 9.32×10^4 for the PIP-NN PES. Obviously, the combination of these two factors for each PES in the overall temperature regime leads to the final forward rate constants as shown in Fig. 5.

In order to perform a deeper analysis of this tunneling behaviour, Fig. 6 shows the minimum energy paths (MPEs) with changes of energy along the reaction path for the three surfaces considered in the present work, PES-2008, PIP-NN, and Δ-ML. Firstly, the Δ-ML surface improves the agreement with the PIP-NN surface with respect to the PES-2008 one, but still does not match it. The PIP-NN is less repulsive and thinner than PES-2008 (and Δ-ML), and consequently, it will provide larger tunneling effects. It must be noted that Fig. 6 is obtained directly from the MEP calculation over each PES, that is, from the saddle point optimized at each level to the reactant and product. This represents a different approach from that used in Fig. 3, where only the PES-2008 MEP was calculated, and then single-point calculations were performed.

(2) Dynamic results

In this section, we focus on the $H + CD_4 \rightarrow HD + CD_3$ reaction using QCT calculations on the three surfaces, PES-2008, PIP-NN, and Δ-ML, at a collision energy of 1.5 eV since most experimental information is available for comparison.

(A) Reaction cross-section, σ_R (Å^2). The QCT results (using both the All and DZPE approaches) at a collision energy of 1.5 eV on the three surfaces are listed in Table 2. Firstly, we observed different behaviours between the two counting

Table 2 Reaction cross section, σ_R (Å^2), at 1.5 eV, for the $H + CD_4$ reaction

Reference	All approach	DZPE approach
PES-2008	0.13	0.03
PIP-NN	0.09	0.01
Δ-ML	0.16	0.05
Experim.[a]	0.14 ± 0.03	

[a] Ref. 56.

methods, All and DZPE approaches, where, as it was expected, the DZPE approach drastically diminishes the reactivity. Secondly, the three surfaces showed similar behaviour. Next, a theory/experiment comparison is performed, which tests both the PES and the dynamics tool. Experimentally, Germann et al.[56] studied this hydrogen abstraction reaction at 1.5 eV by using coherent anti-Stokes Raman spectroscopy (CARS), finding that the total reaction cross section is 0.14 ± 0.03 Å^2, which coincides with most recent measurements,[57,58] and these last ones are performed using relative excitation functions. A note of caution on the last value of Zhang et al.[58] is necessary because, in their original paper (Fig. 1), the ordinate axis presents the values sin Bohr^2 rather than Å^2. We cannot know if this is a typo or not. As compared to the experiment, PES-2008 and Δ-ML reproduce the experimental evidence when the All counting method is considered, but when the DZPE counting method is used, the three surfaces strongly underestimate the experiment.

(B) Product energy partition. Table 3 lists the QCT-DZPE average product fraction of energy in translation, f_{trans}, and internal energy of HD and CD_3 products, $f_{vib}(HD)$, $f_{rot}(HD)$, $f_{vib}(CD_3)$, and $f_{rot}(CD_3)$, at 1.5 eV, for the three surfaces analysed in the present work, together with experimental values[56] for comparison. Firstly, the three surfaces give a similar picture of this dynamic property, with ∼50–60% of energy deposited as translational energy, small energy, ∼7%, deposited as CD_3
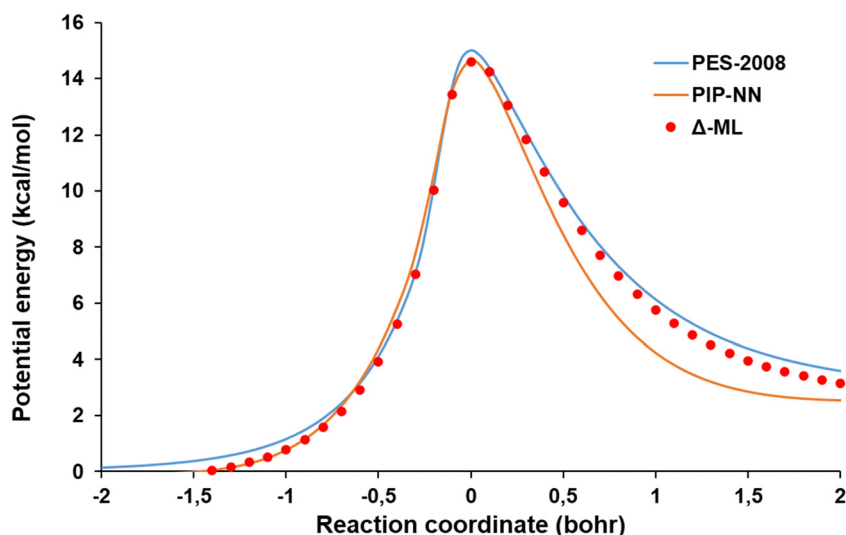
Fig. 6 Minimum energy paths for the three surfaces, PES-2008, PIP-NN, and Δ-ML, are each optimized at their respective level. Note that s = 0 in the reaction coordinate axis corresponds to the respective optimized saddle point at each level.

19210 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19204–19215

This journal is © the Owner Societies 2025

Table 3 Product energy partitioning in percentages (the maximum error bar calculated in this work is ± 1) for the H + CD$_4$ hydrogen abstraction reaction at 1.5 eV

Reference	f_{trans}	f_{vib}(HD)	f_{rot}(HD)	HD$_{\text{internal}}$	f_{vib}(CD$_3$)	f_{rot}(CD$_3$)	CD$_3$, internal
PES-2008	56	17	20	37	5	2	7
PIP-NN	50	24	19	43	4	3	7
Δ-ML	59	10	24	34	4	3	7
Experim.[a]		7	9	16			

[a] Ref. 56.

internal energy, and noticeable energy, ~ 30–40%, deposited as HD-coproduct internal energy. These results show that even PES-2008 surfaces give a reasonable description of this dynamic's magnitude. These theoretical results strongly contrast with the experimental evidence at this same collision energy,[55] where only a small fraction of the available energy appears as HD-coproduct internal energy, 16%, with 7% and 9% as HD vibration and rotation, respectively. This strong theoretical/experimental discrepancy, 30–40% versus 16%, in HD internal energy, could be due to limitations of the QCT approach due to its classical nature, but also to experimental problems. So, Hu et al.[57] suggested that the conclusions from the CARS (coherent anti-Stokes Raman scattering spectroscopy) experimental study[56] might need to be reinterpreted. In any case, the theoretical overestimation of the HD rotational energy as compared with the experiment will have repercussions on the HD rotovibrational distribution, giving a priori, hotter rotational distributions. Later, in Section C, we will return to this issue.

(C) HD product vibrational and rotational distributions. The QCT-DZPE HD product vibrational distributions at a collision energy of 1.5 eV on the three surfaces are listed in Table 4, together with experimental values[56] for comparison. Considering all rotational states, the HD($v = 0 + 1$) vibrational states present similar behaviour for the three surfaces, with populations of 96%, 85%, and 93%, using, respectively, PES-2008, PIP-NN, and Δ-ML surfaces. In addition, they reproduce the experimental evidence, where the HD molecules formed in the $v = 0$ and $v = 1$ vibrational states represent more than 95% of the population.

With respect to the HD product rotational distributions associated with these vibrational states, $v = 0$ and $v = 1$, Fig. 7 plots these distributions using the QCT-DZPE approach at a collision energy of 1.5 eV. Note that the QCT outcome gives a fractional rotational number, and they are truncated to its lower integer part. In general, all three surfaces present similar behaviour, with a normal negative correlation of vibrational and rotational distributions, i.e., hotter vibrational states correlate with colder rotational populations, where the rotational distributions peak in similar j numbers (Table 4). This is the universal behaviour expected for direct bimolecular reactions, but it contrasts with the experimental measures.[56] In order to go deeper into the rotational behaviour, in the previous section (Table 3), we observed that the QCT method tends to give rotational distributions hotter and wider than the experiment, $\sim 20\%$ versus 9%, independent of the potential energy surface used. We suppose that this anomalous behaviour is an artifact of the QCT method due to its classical nature. In addition, as will be analysed in the next section, this overestimation has important consequences in the scattering angle distributions. Using the harmonic approximation, the HD($v = 1$) vibrational state is ~ 11 kcal mol^{-1} above the HD ($v = 0$) vibrational ground state, $v = 3823$ cm^{-1}. However, the QCT results show that the HD product rotational populations are non-negligible up to $j = 14$. Taking into account the HD product rotational constant, $B = 45.65$ cm^{-1}, this implies an energy of ~ 17 kcal mol^{-1} (using the rigid rotor model), which is higher than the energy of the first vibrational state. As a consequence, the j maximum value allowed at the HD($v = 0$) vibrational state is $j = 8$, i.e., six units lower than the QCT results.

(D) HD product angular distributions. At a collision energy of 1.5 eV, Fig. 8 (upper panel) shows the product angular scattering distribution of the CD$_3$ product with respect to the incident H atom for three surfaces, PES-2008, PIPNN, and Δ-ML, by using QCT results. The three surfaces yield predominantly a forward scattered CD$_3$ product, with a small participation of sideways scattering, especially the PIPNN surface, suggesting a rebound mechanism associated with low impact parameters. No experimental data are available at this collision energy, but at 1.2 and 1.9 eV, Zare and coworkers[59–61] measured this angular distribution, finding that the CD$_3$ scattering (with respect to the H atom incident, Fig. 8, lower panel) was mainly sideways and backward, suggesting a stripping mechanism and large impact parameters.

This experimental behaviour strongly contrasts with the theoretical finding. A priori, this experimental/theoretical discrepancy would indicate problems with the potential energy surfaces, which do not reproduce the experiments. However, when one is comparing theory with the experiment, all theoretical tools are being tested, in this case, PES and dynamics tool, QCT. In a previous paper of our group,[40] using the PES-2008 surface,

Table 4 HD(v,j) product rotovibrational populations in percentages for the H + CD$_4$ hydrogen abstraction reaction at 1.5 eV

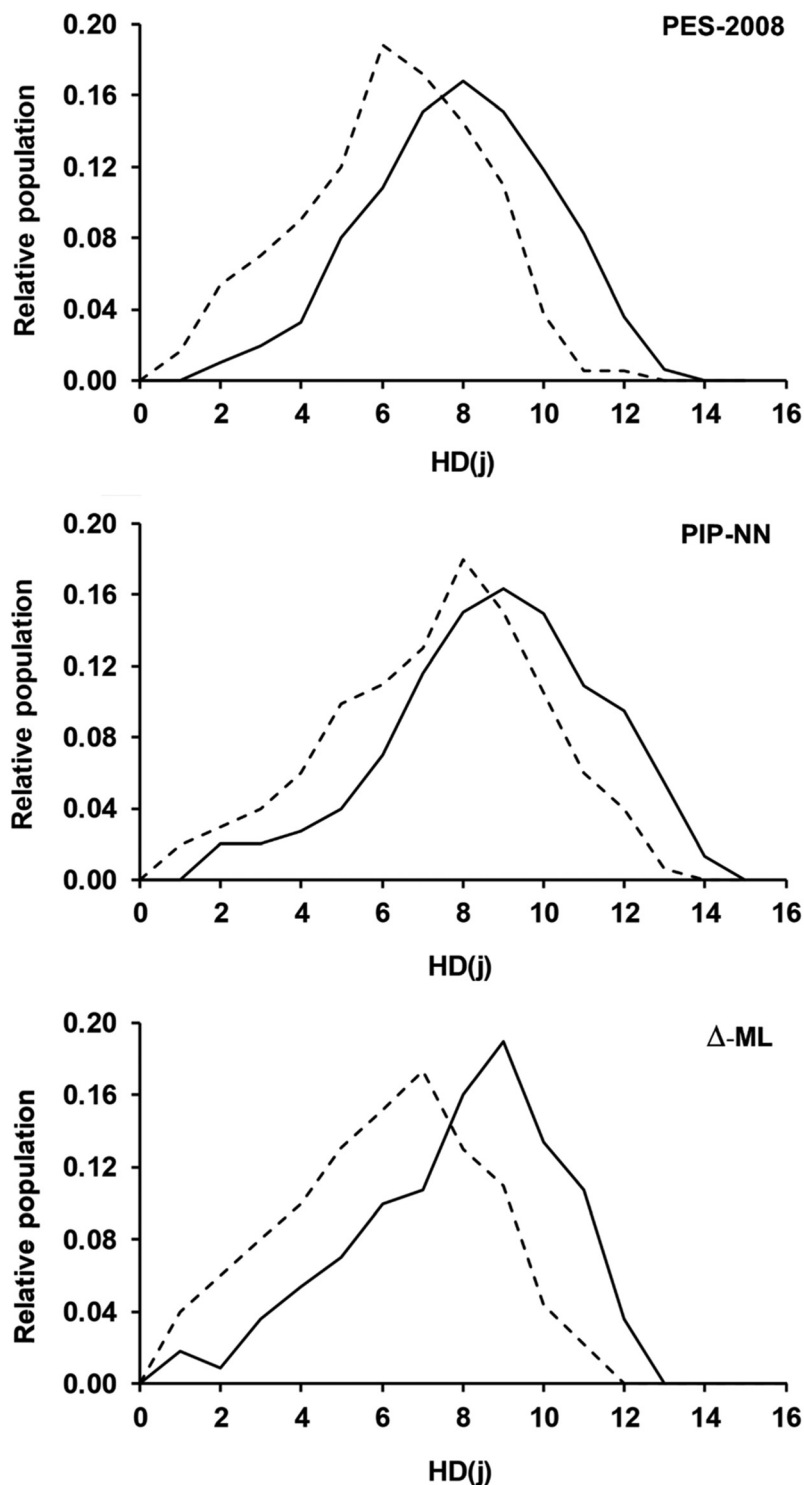| Reference	HD ($v = 0$)	HD ($v = 1$)	HD ($v = 0 + 1$)	HD ($v = 0, j_{\text{max}}$)	HD ($v = 1, j_{\text{max}}$)
PES-2008	60	36	96	8	6
PIP-NN	42	43	85	9	8
Δ-ML	66	27	93	9	7
Experim.[a]			>95		

[a] Ref. 56.

Fig. 7 HD(v, j) product rotational distributions for the three surfaces, PES-2008, PIP-NN, and Δ-ML, at 1.5 eV of collision energy for the H + CD_4 reaction using QCT results. The solid line corresponds to the HD(v = 0) vibrational state and the dashed line corresponds to the HD(v = 1) excited vibrational state.

reduced dimensionality quantum scattering calculations using the rotating line umbrella, RLU, model,[62–64] in brief, QM-RLU/PES-2008, were performed. In this model, only three degrees of freedom are explicitly treated quantum dynamically: the forming H–D stretching vibration, the breaking C–D stretching vibration, and

the umbrella motion. The scattering distribution is also plotted in Fig. 8, lower panel. Now, the experimental evidence is reasonably reproduced, taking into account the experimental uncertainties. In this paper,[40] we also noted that the discrepancies were due to QCT limitations and, more specifically, to the overestimation of
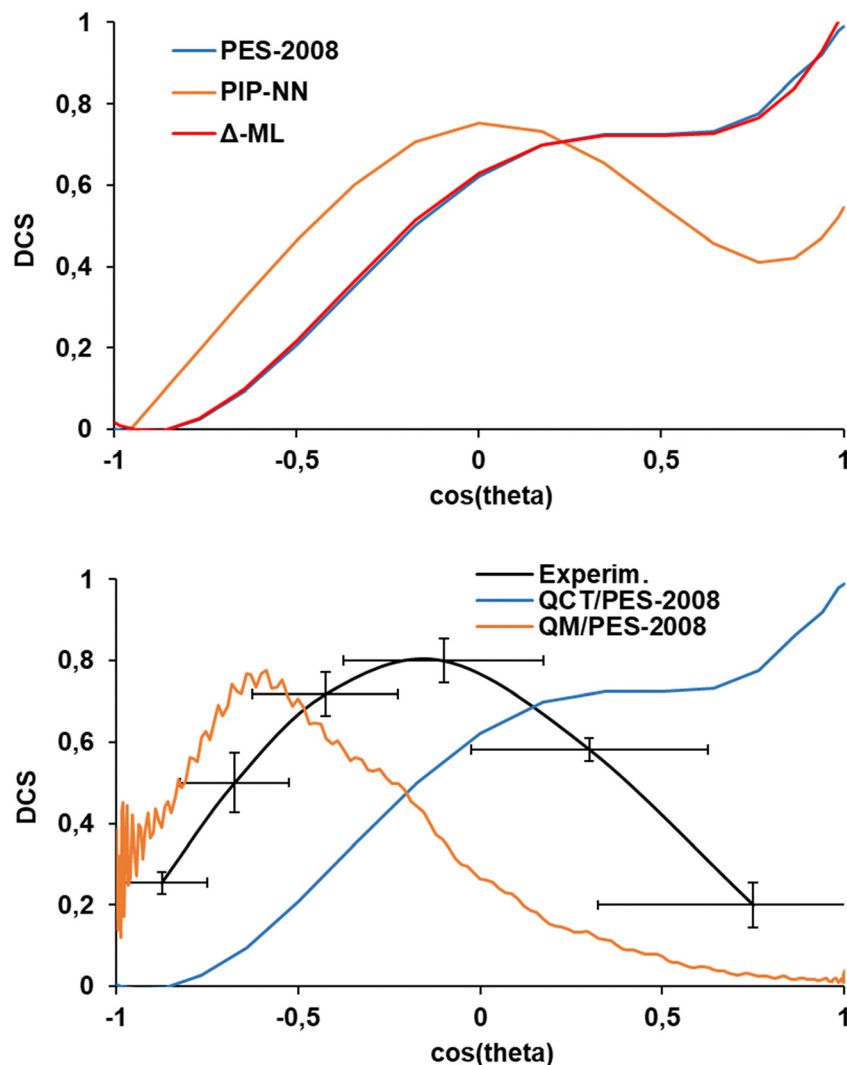
19212 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19204–19215

This journal is © the Owner Societies 2025

Fig. 8 CD$_3$ product angular distribution with respect to the incident H atom for the H + CD$_4$ hydrogen abstraction reaction at 1.5 eV. Upper panel: for the three surfaces, PES-2008, PIP-NN, and Δ-ML, using QCT results. Lower panel: QCT and QM results on the PES-2008 surface[40] and experimental results (at 1.2 eV) from ref. 59–61.

the HD rotational excitation, which yield the scattering distribution observed (Fig. 8, upper panel).

Conclusions

In the present work, we apply the ANN technique in Δ-ML for developing reactive surfaces, in which a correction PES was developed to bring the LL-PES to the HL-PES based on a small number of configurations. This ANN-based Δ-ML approach is similar to previously proposed Δ-ML methods. This strategy has been applied to the well-known polyatomic H + CH$_4$/CD$_4$ hydrogen abstraction reactions. The main difference with respect to other Δ-ML methods is the use of analytical surfaces previously developed for this reactive system. Thus, as LL, we used the information from our analytical full-dimensional PES-2008 and the HL information proceeds from the PIP-NN surface of Li and coworkers (2015),[43] which is the most accurate surface for this polyatomic

system. The efficacy of the present model can be evaluated by analysing the RMSE (root mean square error) values, which decrease from 0.5447 kcal mol^{-1} using the difference between V_{HL} and V_{LL} energies to 0.0263 kcal mol^{-1} after implementing the correction with the Δ-ML model. In addition, the quality of the Δ-ML strategy is demonstrated by performing kinetics and dynamics calculations. Kinetically, VTST/MT theory is used on the three surfaces, PES-2008, PIP-NN, and Δ-ML, finding that the rate constants obtained with the Δ-ML model agree with those obtained with the HL-PES, improving the description of the PES-2008, especially at low temperatures, where quantum tunneling is more important. Dynamically, QCT calculations were performed on the three surfaces for the H + CD$_4$ reaction, studying dynamic properties, such as the reaction cross section, HD product roto-vibrational distributions, and product angular distributions. These kinetic and dynamic results suggest the promise of the Δ-ML alternative in developing highly accurate surfaces for large reactive systems.

Conflicts of interest

The authors have no conflicts to disclose.

Data availability

The potential energy surfaces used in this study have been previously published (see ref. 39 and 42). The present Δ-ML PES is constructed directly based on these established surfaces.

Acknowledgements

References

1 J. Behler, *J. Chem. Phys.*, 2016, **145**, 170901see also references therein.

2 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**(14), 146401.

3 N. Artrith and A. M. Kolpak, *Comput. Mater. Sci.*, 2015, **110**, 20–28.

4 J. Behler, *Phys. Chem. Chem. Phys.*, 2011, **13**(40), 17930–17955.

5 J. Behler, *J. Phys.: Condens. Matter*, 2014, **26**(18), 183001.

6 J. R. Boes, M. C. Groenenboom, J. A. Keith and J. R. Kitchin, *Int. J. Quantum Chem.*, 2016, **116**(13), 979–987.

7 P. E. Dolgirev, I. A. Kruglov and A. R. Oganov, *AIP Adv.*, 2016, **6**(8), 085318.

8 M. Gastegger and P. Marquetand, *J. Chem. Theory Comput.*, 2015, **11**(5), 2187–2198.

9 S. Manzhos, R. Dawes and T. Carrington, *Int. J. Quantum Chem.*, 2015, **115**(16), 1012–1020.

10 S. K. Natarajan, T. Morawietz and J. Behler, *Phys. Chem. Chem. Phys.*, 2015, **17**(13), 8356–8371.

11 N. Lubbers, J. S. Smith and K. Barros, *J. Chem. Phys.*, 2018, **148**(24), 241715.

12 J. S. Smith, O. Isayev and A. E. Roitberg, *Chem. Sci.*, 2017, **8**(4), 3192–3203.

13 B. Kolb, L. C. Lentz and A. M. Kolpak, *Sci. Rep.*, 2017, 7, 1192.

14 K. Schutt, P. J. Kindermans, H. E. S. Felix, S. Chmiela, A. Tkatchenko and K. R. Muller, *Advances in Neural Information Processing Systems*, 2017, pp. 992–1002.

15 L. Zhang, J. Han, H. Wang, R. Car and E. Weinan, *Phys. Rev. Lett.*, 2018, **120**(14), 143001.

16 K. Ryczko, K. Mills, I. Luchak, C. Homenick and I. Tamblyn, *Comput. Mater. Sci.*, 2018, **149**, 134–142.

17 K. Yao, J. E. Herr, D. W. Toth, R. Mckintyre and J. Parkhill, *Chem. Sci.*, 2018, **9**(8), 2261–2269.

18 A. P. Bartok, M. C. Payne, R. Kondor and G. Csanyi, *Phys. Rev. Lett.*, 2010, **104**(13), 136403.

19 W. J. Szlachta, A. P. Bartok and G. Csanyi, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **90**(10), 104108.

20 V. L. Deringer and G. Csanyi, *Phys. Rev. B*, 2017, **95**, 094203.

21 V. L. Deringer, C. J. Pickard and G. Csanyi, *Phys. Rev. Lett.*, 2018, **120**(15), 156001.

22 A. Grisafi, D. M. Wilkins, G. Csanyi and M. Ceriotti, *Phys. Rev. Lett.*, 2018, **120**(3), 036002.

23 C. Qu, Q. Yu and J. M. Bowman, *Annu. Rev. Phys. Chem.*, 2018, **69**, 151–175.

24 B. J. Braams and J. M. Bowman, *Int. Rev. Phys. Chem.*, 2009, **28**, 577–606.

25 J. Zuo, Y. Li, H. Guo and D. Xie, *J. Phys. Chem. A*, 2016, **120**(20), 3433–3440.

26 M. Bai, D. Lu, Y. Li and J. Li, *Phys. Chem. Chem. Phys.*, 2016, **18**, 32031–32041.

27 J. Zuo, C. Xie, H. Guo and D. Xie, *J. Phys. Chem. Lett.*, 2017, **8**(14), 3392–3397.

28 J. Zhu, V. Q. Vuong, B. G. Sumpter and S. Irle, *MRS Commun.*, 2019, **9**, 846–859, DOI: **10.1557/mrc.2019.80**.

29 P. O. Dral, A. Owens, A. Dral and G. Csanyi, *J. Chem. Phys.*, 2020, **152**, 204110.

30 A. Nandi, C. Qu, P. L. Houston, R. Conte and J. M. Bowman, *J. Chem. Phys.*, 2021, **154**, 051102.

31 C. Qu, P. L. Houston, R. Conte, A. Nandi and J. M. Bowman, *J. Phys. Chem. Lett.*, 2021, **12**, 4902–4909.

32 Y. Liu and J. Li, *J. Phys. Chem. Lett.*, 2022, **13**, 4729–4738.

33 R. Ramakrishnan, P. O. Dral, M. Rupp and O. A. von Lilienfeld, *J. Chem. Theory Comput.*, 2015, **11**, 2087–2096.

34 B. Fu, X. Xu and D. H. Zhang, *J. Chem. Phys.*, 2008, **129**, 011103.

35 W.-P. Hu, Y.-P. Liu and D. G. Truhlar, *J. Chem. Soc., Faraday Trans.*, 1994, **90**, 1715–1725.

36 D. K. Malick, G. A. Petersson and J. A. Montgomery, Jr, *J. Chem. Phys.*, 1998, **108**, 5704–5713.

37 M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357.

38 J. Espinosa-Garcia and J. C. Corchado, *J. Chem. Phys.*, 1996, **100**, 16561.

39 J. Espinosa-Garcia, *J. Chem. Phys.*, 2002, **116**, 10664.

40 J. C. Corchado, J. L. Bravo and J. Espinosa-Garcia, *J. Chem. Phys.*, 2009, **130**, 184314.

41 X. Zhang, B. J. Braams and J. M. Bowman, *J. Chem. Phys.*, 2006, **124**, 021104.

42 Y. Zhou, B. Fu, C. Wang, M. A. Collins and D. H. Zhang, *J. Chem. Phys.*, 2011, **134**, 064323.

19214 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19204–19215

This journal is © the Owner Societies 2025

43 J. Li, J. Chen, Z. Zhao, D. Xie, D. H. Zhang and H. Guo, *J. Chem. Phys.*, 2015, **142**, 204302.

44 T. Wu, H.-J. Werner and U. Manthe, *J. Chem. Phys.*, 2006, **124**, 164307.

45 J. Pu and D. G. Truhlar, *J. Chem. Phys.*, 2002, **116**, 1468.

46 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, Scikit-learn: Machine Learning in Python, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

47 D. L. Baulch, C. J. Cobos, R. A. Cox, C. Esser, P. Frank, Th Just, J. A. Kerr, M. J. Pilling, J. Troe, R. W. Walker and J. Warnatz, *J. Phys. Chem. Ref. Data*, 2005, **34**(3), 757–1397.

48 KIDA2, Kinetic Database for Astrochemistry.

49 B. C. Garrett and D. G. Truhlar, *J. Am. Chem. Soc.*, 1979, **101**, 4534.

50 D. G. Truhlar, A. D. Isaacson and B. C. Garrett, Generalized Transition State Theory, in *The Theory of Chemical Reactions*, ed. M. Baer, CRC, Boca Raton, FL, 1985, vol. 4.

51 Y. P. Liu, D. H. Lu, A. Gonzalez-Lafont, D. G. Truhlar and B. C. Garrett, *J. Am. Chem. Soc.*, 1993, **115**, 7806–7817.

52 Y. Y. Chuang and D. G. Truhlar, *J. Phys. Chem. A*, 1997, **101**, 3808–3814.

53 J. Zheng, J. L. Bao, R. Meana-Pañeda, S. Zhang, B. J. Lynch, J. C. Corchado, Y. Y. Chuang, P. L. Fast, W. P. Hu, Y. P. Liu, G. C. Lynch, K. A. Nguyen, C. F. Jackels, A. Fernandez Ramos, B. A. Ellingson, V. S. Melissas, J. Villa, I. Rossi, E. L. Coitiño, J. Pu, T. V. Albu, A. Ratkiewicz, R. Steckler, B. C. Garrett, A. D. Isaacson and D. G. Truhlar, *Polyrate-version 2017-C*, University of Minnesota, Minneapolis, MN, 2017.

54 X. Hu, W. L. Hase and T. Pirraglia, *J. Comput. Chem.*, 1991, **12**, 1014.

55 W. L. Hase, R. J. Duchovic, X. Hu, A. Komornicki, K. F. Lim, Dh Lu, G. H. Peslherbe, K. N. Swamy, S. R. van de Linde, A. J. C. Varandas, H. Wang and R. J. Wolf, VENUS96: A General Chemical Dynamics Computer Program, *QCPE Bull.*, 1996, **16**, 43.

56 G. Germann, Y. Huh and J. Valentini, *J. Chem. Phys.*, 1992, **96**, 1957.

57 W. Hu, G. Lendvay, D. Troya, G. C. Schatz, J. P. Camden, H. A. Bechtel, D. J. A. Brown, M. R. Martin and R. N. Zare, *J. Phys. Chem. A*, 2006, **110**, 3017.

58 W. Zhang, Y. Zhou, G. Wu, Y. Lu, H. Pan, B. Fu, Q. Shuai, L. Liu, S. Liu, L. Zhang, B. Jiang, D. Dai, S. Lee, Z. Xie, B. J. Braams, J. M. Bowman, M. A. Collins, D. H. Zhang and X. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 12782.

59 J. P. Camden, H. A. Bechtel and R. N. Zare, *Angew. Chem., Int. Ed.*, 2003, **42**, 5227.

60 J. P. Camden, H. A. Bechtel, D. J. A. Brown, M. R. Martin, R. N. Zare, W. Hu, G. Lendvay, D. Troya and G. C. Schatz, *J. Am. Chem. Soc.*, 2005, **127**, 11898.

61 J. P. Camden, W. Hu, H. A. Bechtel, D. J. A. Brown, M. R. Martin, R. N. Zare, G. Lendvay, D. Troya and G. C. Schatz, *J. Phys. Chem. A*, 2006, **110**, 677.

62 H.-G. Yu and G. Nyman, *Chem. Phys. Lett.*, 1998, **298**, 27.

63 H.-G. Yu and G. Nyman, *Phys. Chem. Chem. Phys.*, 1999, **1**, 1181.

64 H.-G. Yu and G. Nyman, *J. Chem. Phys.*, 1999, **111**, 6693.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19204–19215 | **19215**