New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

FAST TRACK COMMUNICATION • OPEN ACCESS

Bayesian optimization for the inverse scattering problem in quantum reaction dynamics

To cite this article: R A Vargas-Hernández *et al* 2019 *New J. Phys.* **21** 022001

View the article online for updates and enhancements.

You may also like

- Significant non-adiabatic effects of the $K(4s^2S) + H_2$ reaction
Wentao Li, Li Wen, Xianghong Niu et al.
- Leveraging trust for joint multi-objective and multi-fidelity optimization
Faran Irshad, Stefan Karsch and Andreas Döpp
- A review of dynamical resonances in A + BC chemical reactions
Zefeng Ren, Zhigang Sun, Donghui Zhang et al.

This content was downloaded from IP address 18.227.0.192 on 02/05/2024 at 16:06

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
8 October 2018

ACCEPTED FOR PUBLICATION
22 January 2019

PUBLISHED
28 February 2019

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

FAST TRACK COMMUNICATION

Bayesian optimization for the inverse scattering problem in quantum reaction dynamics

R A Vargas-Hernández[1], **Y Guan**[2,3], **D H Zhang**[2,3] **and R V Krems**[1]

[1] Department of Chemistry, University of British Columbia, Vancouver, BC V6T 1Z1, Canada
[2] State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian 116023, People's Republic of China
[3] University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

E-mail: rkrems@chem.ubc.ca

Keywords: bayesian optimization, gaussian processes, inverse scattering problem

Abstract

We propose a machine-learning approach based on Bayesian optimization to build global potential energy surfaces (PES) for reactive molecular systems using feedback from quantum scattering calculations. The method is designed to correct for the uncertainties of quantum chemistry calculations and yield potentials that reproduce accurately the reaction probabilities in a wide range of energies. These surfaces are obtained automatically and do not require manual fitting of the *ab initio* energies with analytical functions. The PES are built from a small number of *ab initio* points by an iterative process that incrementally samples the most relevant parts of the configuration space. Using the dynamical results of previous authors as targets, we show that such feedback loops produce accurate global PES with 30 *ab initio* energies for the three-dimensional $H + H_2 \rightarrow H_2 + H$ reaction and 290 *ab inito* energies for the six-dimensional $OH + H_2 \rightarrow H_2O + H$ reaction. These surfaces are obtained from 360 scattering calculations for H_3 and 600 scattering calculations for OH_3. We also introduce a method that quickly converges to an accurate PES without the *a priori* knowledge of the dynamical results. By construction, our method illustrates the lowest number of potential energy points (i.e. the minimum information) required for the non-parametric construction of global PES for quantum reactive scattering calculations.

1. Introduction

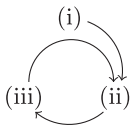
The accurate description of physical processes involving microscopic scattering of molecules is hampered by the lack of knowledge of accurate potential energy surfaces (PES) underlying the scattering events. While sophisticated experiments probing the outcome of molecule—surface scattering or molecule—molecule collisions can be designed, the theoretical description of such experiments is limited by the difficulty of the electronic structure calculations necessary to compute the microscopic scattering matrices. Instead of computing the PES for molecular scattering from first principles, one can obtain 'empirical' PES constructed based on the information about the experimental data and designed to yield an accurate description of the experimental measurements. In the present work, we refer to the construction of such empirical PES as the inverse scattering problem. In general, the inverse scattering problem is a complex task due to the many degrees of freedom relevant for molecular interactions and the complexity of the scattering event. Here, we propose a machine-learning (ML) method based on Bayesian optimization (BO) for the inverse scattering problem. While the approach is general and can—in principle—be applied to any problem with a PES $\leftrightarrow$ observable correlation, we discuss the application of the method to building global PES for chemically reactive molecular systems based on feedback from quantum scattering calculations.

Any quantum dynamics calculation of molecular collision observables involves three steps: (i) computing the potential energy for a wide range of relative atomic coordinates by an *ab initio* quantum chemistry method;

© 2019 The Author(s). Published by IOP Publishing Ltd on behalf of the Institute of Physics and Deutsche Physikalische Gesellschaft

(ii) fitting these energy points to construct a PES; (iii) integrating the Schrödinger equation for the motion of the atomic nuclei on this PES. Unfortunately, it is impossible to compute the potential energy in step (i) without errors and any theoretical predictions of observables are subject to uncertainties stemming from the errors of quantum chemistry calculations. These uncertainties have a particularly large effect on reactions at low temperatures and make quantitative predictions of reaction probabilities at ultracold temperatures impossible [1]. Therefore, it is necessary to develop approaches that correct the errors of the *ab initio* calculations and produce PES yielding the exact match between quantum reaction dynamics calculations and experimental measurements.

For chemically non-reactive two- or three-atom systems, it was previously shown that empirical PES can be derived from the experimental data [2–6]. These approaches generally involve an iterative feedback loop that can be schematically illustrated as follows:

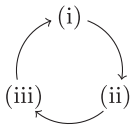


(1)

The PES is computed in step (i), an analytical fit of the PES is generated in step (ii) and the fit is then morphed through a feedback loop involving a series of quantum dynamics calculations.

Such an approach is impossible to apply to chemically reactive systems, especially ones involving more than three atoms, for three reasons. First, the active configuration space describing reactive systems is more complex, involving multiple reaction channels, and it is not *a priori* known which part of the configuration space is most important for the outcome of a reactive process. Second, each reaction dynamics calculation is time consuming. Third, fitting PES for reactive systems is a complex task that almost always requires manual work [7–11].

Here, we design a method that makes possible the following optimization loop for three-atom (3D) and four-atom (6D) chemically reactive systems:



(2)

This approach allows one to compute and add *ab initio* points to the PES incrementally at each iteration. This has four major advantages over conventional approaches to constructing PES:

- First, the *ab initio* points are only placed in the parts of the configuration space most relevant for the dynamics. The final PES thus offers unique information on the parts of the PES that determine the specific reaction features under study.
- Second, this approach eliminates the need for a large number of the *ab initio* points, reducing the computational effort associated with quantum chemistry calculations.
- Third, as explained below, this approach eliminates the laborious task of fitting the PES. The PES is produced automatically as a mean of a multi-variate distribution.
- Finally, the PES thus constructed is guaranteed to yield quantum reaction dynamics observables that agree with the experimental data.

There are two key steps introduced here that make the above optimization loop possible. The reaction observables are approximated by a ML model, which is a function of another ML model describing the PES. In the present approach, both of these ML models are provided by Gaussian process (GP) regression [21]. GP regression is a statistical learning technique, which provides a prediction and an uncertainty of the prediction. As described below, the two-tiered ML model is used here to eliminate the need for fitting the PES manually and the uncertainty of the GP prediction is used here to make the optimization loops extremely efficient by means of BO.

The remainder of the article is organized as follows. We begin by reviewing the application of GP regression to the construction of PES. The following section describes the main ideas behind BO based on GP regression and how BO can be applied to construct the PES based on the information about the scattering observables. The subsequent sections present the results demonstrating the efficiency of the BO approach.

2. GP regression for PES

In our approach, the global reactive PES as a function of the multi-dimensional vector $\boldsymbol{r}$ is given as

$$V(\boldsymbol{r}) = \sum_{i=1}^{n} w_i(\boldsymbol{r}) E_i, \tag{3}$$

where E_i are the *ab initio* energy points and the weights (w_i) are determined by the two-tiered GP model to yield the best outcome of the quantum scattering calculation. GPs have been previously used for interpolating PES for molecular dynamics applications [12–16], spectroscopic line calculations [17, 18] and molecular scattering calculations [19, 20]. We emphasize that equation (3) is not a fit of the PES but a non-parametric regression. The coefficients w_i are chosen such that equation (3) passes through the potential energy points. Between the points, equation (3) represents a mean of a distribution that changes in response to the addition of an energy point to the ensemble $\{E_i\}$ of *ab initio* points. More specifically, the weights in equation (3) are the elements of the following vector [21]:

$$w_i(\boldsymbol{r}_0) = [\mathbf{A}_0^{\top}(\boldsymbol{r}_0)\mathbf{A}^{-1}]_i, \tag{4}$$

where $\mathbf{A}$ is an $n \times n$ matrix of the covariances between all pairs of E_i and $\mathbf{A}_0$ is a vector of covariances between the value of energy at $\boldsymbol{r}_0$ and all known E_i. The procedure to obtain $\mathbf{A}$ and $\mathbf{A}_0$ is described elsewhere [17, 22] and is briefly summarized in the remainder of this section.

GP regression is a non-parametric supervised ML algorithm [21], which belongs to a class of kernel regression algorithms. The prediction of a GP is a Gaussian distribution. Essentially, we seek to answer the following question: given the distribution of *n ab initio* energy points E_i located at $\boldsymbol{r}_i$ and collectively represented by vector $\mathbf{y}$, what is the value of the potential energy at any point $\boldsymbol{r}$ of the configuration space? GP regression produces a probability distribution for this value. This conditional distribution is a Gaussian with the conditional mean μ and conditional variance σ derived [21] to have the following form:

$$\mu(\boldsymbol{r}_0) = \beta + \mathbf{A_0}^{\top}\mathbf{A^{-1}}(\mathbf{y} - \beta), \tag{5}$$

$$\sigma(\boldsymbol{r}_0) = \sigma_u(1 - \mathbf{A_0}^{\top}\mathbf{A^{-1}}\mathbf{A_0}), \tag{6}$$

where β and σ_u are the unconditional means of the GP. In the present work, we set β to zero, assuming that we have no prior knowledge of the PES. This is the simplest and least efficient implementation. Here, we use the mean of the conditional distribution (5) as the value for the interpolation.

By construction

$$\mu(\boldsymbol{r}_i) = E_i \tag{7}$$

and

$$\sigma(\boldsymbol{r}_i) = 0, \tag{8}$$

i.e. the Gaussian distributions collapse to a single value at the training points in $\mathbf{y}$ so the function $\mu(\boldsymbol{r})$ passes through the points E_i. It is thus an interpolation technique.

In equations (5) and (6), $\mathbf{A}$ is a square matrix of the covariances between all pairs of energy points $\boldsymbol{E}_i$ in the vector $\mathbf{y}$:

$$\mathbf{A} = \text{const} \times \begin{pmatrix} 1 & R(\boldsymbol{r}_1, \boldsymbol{r}_2) & \cdots & R(\boldsymbol{r}_1, \boldsymbol{r}_i) \\ R(\boldsymbol{r}_2, \boldsymbol{r}_1) & 1 & & \vdots \\ \vdots & & \ddots & \\ R(\boldsymbol{r}_i, \boldsymbol{r}_1) & \cdots & & 1 \end{pmatrix} \tag{9}$$

and $\mathbf{A}_0$ is a vector of covariances between the value of the potential energy at $\boldsymbol{r}_0$ and all the values E_i in the vector $\mathbf{y}$:

$$\mathbf{A}_0 = \begin{pmatrix} R(\boldsymbol{r}_1, \boldsymbol{r}_0) \\ R(\boldsymbol{r}_2, \boldsymbol{r}_0) \\ \cdots \\ R(\boldsymbol{r}_i, \boldsymbol{r}_0) \end{pmatrix}. \tag{10}$$

In order to predict the distribution of possible values of energy at $\boldsymbol{r}_0$, the method thus relies on the entire vector of known values of energy in the entire configuration space. Clearly, the effect of the energy points farther away from $\boldsymbol{r}_0$ is smaller than the effect of the energy points nearby. This is accounted for by the mathematical form of the correlation functions $R(\boldsymbol{r}_i, \boldsymbol{r}_j)$, which must decay as the distance between $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ increases.

The GP model is trained by determining the best covariance matrix $\mathbf{A}$ for a given set of values in $\mathbf{y}$. To do this, the functions $R(\boldsymbol{r}_i, \boldsymbol{r}_j)$ are approximated by a simple analytical function. For this work we used the *Matérn* function

$$R(\boldsymbol{r}_i, \boldsymbol{r}_j) = \left(1 + \sqrt{5}\, r(\boldsymbol{r}_i, \boldsymbol{r}_j) + \frac{5}{3} r^2(\boldsymbol{r}_i, \boldsymbol{r}_j)\right) \times \exp\left(-\sqrt{5}\, r^2(\boldsymbol{r}_i, \boldsymbol{r}_j)\right) \tag{11}$$

with

$$r^2(\boldsymbol{r}_i, \boldsymbol{r}_j) = (\boldsymbol{r}_i - \boldsymbol{r}_j)^\top \boldsymbol{M} (\boldsymbol{r}_i - \boldsymbol{r}_j)$$
$$= \left(r_i^1 - r_j^1, \ \cdots, \ r_i^d - r_j^d\right) \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_d \end{pmatrix} \begin{pmatrix} r_i^1 - r_j^1 \\ \vdots \\ r_i^d - r_j^d \end{pmatrix},$$

where d is the number of dimensions and the diagonal matrix $\boldsymbol{M}$ contains the kernel parameters θ_i. To find the best *estimates* of the kernel parameters we maximize the log *marginal likelihood* with respect to the parameters. The log marginal likelihood is

$$\log p(\mathbf{y}|\boldsymbol{r}, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^\top \mathbf{A}^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{A}| - \frac{n}{2}\log(2\pi), \tag{12}$$

where $|\mathbf{A}|$ is the determinant of the matrix $\mathbf{A}$ and n is the number of training points.

Note that this procedure does not produce analytical fits of the values in $\mathbf{y}$. Neither should the best covariance function R found by maximizing the log-likelihood function be considered an analytical fit of the covariances. Rather, R is an estimator of the covariances. The function R found by training the GP model provides a function, which parametrizes the GP predictive distributions. Although the efficiency and the properties (such as the differentiability) of the resulting GP model depend on the choice of the mathematical function for R, this choice is not unique and should lead to the same results for large n. A simple Gaussian function or an exponentially decaying function could have been chosen in place of the *Matérn* function, leading to the same result, although with a larger n (for the present applications).

3. BO of scattering observables

In the previous section, we described how GP regression can be used for step (ii) in equation (2). Using GP models for step (ii) provides a way to automatically construct the PES that can be used for quantum dynamics calculations in step (iii). Because GP regression produces the PES in the form of equation (3), this automates the optimization cycle (2), i.e. a different global PES is produced by GP regression for each loop of equation (2) without manual work. However, each feedback loop in equation (2) involves a scattering calculation, which is often time-consuming. A typical scattering calculation of the reaction observables takes minutes to hours of CPU time, depending on the complexity of the reaction system. Therefore, in general, the feedback loop (2) is considered unfeasible. In this work we show that such feedback loops can be made very efficient (converging to an accurate surface very quickly) and hence feasible by means of BO.

In order to apply BO, we need to construct a machine learning model of the *quantum dynamics results*. We do this also by GP regression. The GP model of quantum dynamics results is trained exactly as described in the previous section, with the training points being the results of quantum dynamics calculations instead of the *ab initio* energy points. This produces a GP model $\mathcal{F}[\mathcal{G}]$ of the scattering calculation results, which is a function of another GP model $\mathcal{G}$ giving the PES. Once $\mathcal{F}[\mathcal{G}]$ is trained, one can apply BO to find the optimal surface $\mathcal{G}$.

The goal of BO is to find the global minimizer (or maximizer) of an unknown objective function f

$$\boldsymbol{x}_* = \arg\max\left[f(\boldsymbol{x})\right] = \arg\min[-f(\boldsymbol{x})]. \tag{13}$$

BO is a sequential model-based approach originally designed for unknown objective functions for which: (i) the gradient of $f(\boldsymbol{x})$ is not available or difficult to evaluate and (ii) the cost of evaluating f at $\boldsymbol{x}$ is high [32, 33]. It is ideally suited for the present application as, in our work, the function f represents the results of quantum scattering calculations, which are generally difficult to compute.

Two ingredients are needed in order to build a BO algorithm. First, we need an emulator function $g(\boldsymbol{x})$ that mimics the unknown objective function $f(\boldsymbol{x})$. In our work, this emulator function will be provided by GP model described in the previous section, but now applied to emulate the quantum dynamics results. Second, it is necessary to define an acquisition function $\alpha(\boldsymbol{x})$ that determines the policy evaluation of $f(\boldsymbol{x})$. In other words, this function will direct the algorithm to the part of the parameter space where the original function must be evaluated.

The general recipe for the BO algorithm starts with finding the minimum/maximum of the acquisition function ($\boldsymbol{x}_* = \arg\max \alpha(\boldsymbol{x})$). The function f is then evaluated at this position $f(\boldsymbol{x}_*)$ producing an update to the emulator g. This procedure is repeated iteratively until the minimizer/maximizer of $g(\boldsymbol{x})$ is converged [32, 33].

In the present work, we use the following acquisition function

$$\alpha(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \kappa\sigma(\boldsymbol{x}), \tag{14}$$

where $\mu(\boldsymbol{x})$ and $\sigma(\boldsymbol{x})$ are the mean and the standard deviation of the GP $\mathcal{F}$. The hyperparameter κ balances the trade-off between exploration and exploitation. When $\kappa \gg 1$, the acquisition function stimulates exploration of the global multi-dimensional space because $\alpha(\boldsymbol{x}) \approx \sigma(\boldsymbol{x})$ so the function directs the evaluation of the original function at parts of the space where the GP is the least certain. On the other hand, when $\kappa \ll 1$ and/or when the uncertainty of the GP σ becomes very small, the acquisition function follows more closely the mean of the GP model, which results in a more accurate determination of the extremum.

In this work, we want the algorithm to converge quickly to a substantial (not necessarily global) maximum. Therefore, the value of κ must be chosen to be $\ll 1$. The actual value of κ determines the efficiency of the algorithm. We choose the value of κ by examining the first optimization loop, to ensure that the BO algorithm converges fast and the improved results can be obtained with less than $15 \times \mathcal{N}$ calculations for each iteration, where $\mathcal{N}$ is the dimensionality of the configuration space. The effect of the value of κ on the convergence speed of BO is discussed in a following section.

4. Application to quantum reaction dynamics

We consider two chemical reactions:

$$H + H_2 \rightarrow H_2 + H, \tag{15}$$

$$OH + H_2 \rightarrow H_2O + H. \tag{16}$$

We compute the total reaction probabilities for the reactant molecules in the ground ro-vibrational state and the total angular momentum $J = 0$ using the time-dependent wave packet dynamics approach described in [23–26], explicitly accounting for all degrees of freedom. The basis sets of the reaction dynamics calculations are chosen to ensure full convergence.

Both of these chemical reactions have been studied before [27, 28]. For reaction (5), Su *et al* [27] computed the reaction probabilities with the 3D PES from [29], constructed using an analytical fit to 8701 *ab initio* energy points. For reaction (6), Chen *et al* [28] computed the reaction probabilities with the 6D PES constructed using Neural Network fits to ~17 000 *ab initio* calculations.

In conventional approaches, the global PES is constructed before the dynamical calculations. The present approach is conceptually different. We begin by randomly selecting a small number of points ($n = 20$ for the 3D surface and $n = 280$ for the 6D surface) from the original PESs [28, 29] and construct an approximate PES by GP regression, giving equation (3) as described in [17]. We denote this GP model of the surface by $\mathcal{G}(n)$. This model is likely to be highly inaccurate.

Given $\mathcal{G}(n)$, we ask: if one *ab initio* point is added to the original sample of few points, where in the configuration space should it be added to result in the maximum improvement of the *quantum dynamics results*? This formulation corresponds to a class of reinforcement learning strategies in ML [30, 31]. In principle, this question could be answered without ML by a series of quantum dynamics calculations based on $\mathcal{G}(n+1)$ with the added point moved on a grid in the configuration space. However, in practice, such an approach would be completely unfeasible as it would require about $10^{\mathcal{N}}$ dynamical calculations for each added *ab initio* point, where $\mathcal{N}$ is the dimensionality of the configuration space.

To overcome this problem, we introduce another GP, hereafter denoted by $\mathcal{F}$, representing the reaction probabilities as functions of the location of the added *ab initio* point. This GP model is trained by $15 \times \mathcal{N}$ quantum dynamics calculations. We thus obtain the explicit function of the reaction probabilities $\mathcal{F}[\mathcal{G}(n+1, \boldsymbol{r})]$, where $\boldsymbol{r}$ is the $\mathcal{N}$-dimensional vector denoting the position of the added *ab initio* point in the configuration space. By training $\mathcal{F}$, we also obtain $\Delta\mathcal{F}[\mathcal{G}(n+1, \boldsymbol{r})]$, representing the uncertainty of $\mathcal{F}$ at $\boldsymbol{r}$, as the conditional standard deviation of the GP [17]. The GP $\mathcal{F}$ and the uncertainty $\Delta\mathcal{F}[\mathcal{G}(n+1, \boldsymbol{r})]$ allow us to implement BO as described in the previous section.

We next propose two approaches:

(1) *Fitting known quantum dynamics results.* If the reaction probabilities are known, either from a calculation with the full surface or from an experiment, one can minimize the difference between $\mathcal{F}[\mathcal{G}(n+1, \boldsymbol{r})]$ and the known results, to obtain the optimal value of $\boldsymbol{r}$. This is the position in the configuration space, where the new *ab inito* point must be added. This minimization can then be iterated by setting $n + 1 \Rightarrow n$, until n becomes large enough to produce an accurate surface $\mathcal{G}$.

(2) *Obtaining the best surface without known dynamical results.* If accurate quantum dynamical results are not available, we propose to *maximize* the difference of the reaction probabilities computed in two successive iterations:

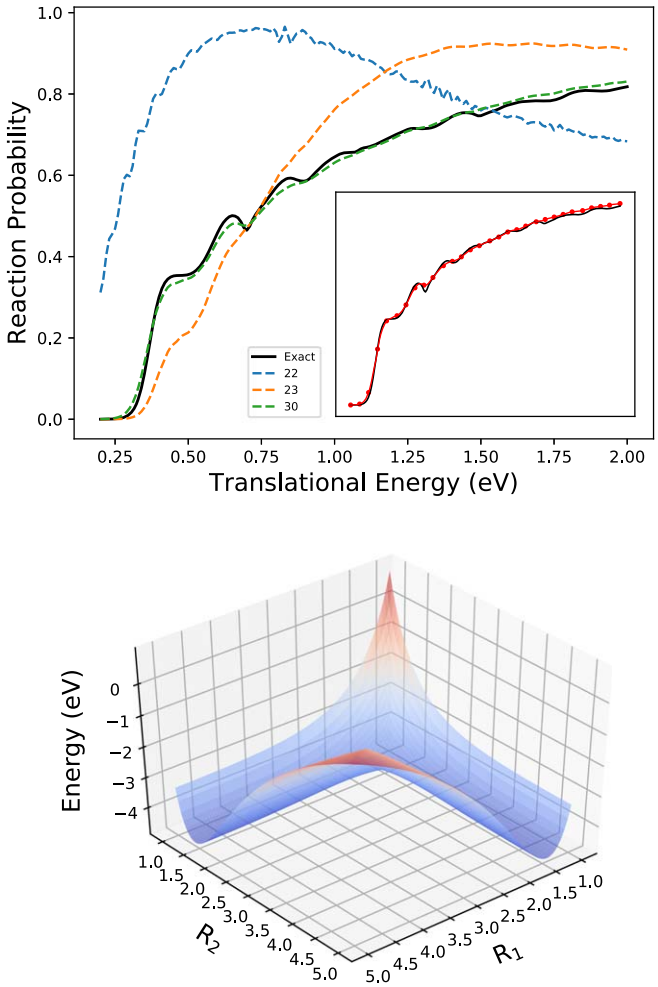


Figure 1. Upper panel: the reaction probability for the $H_2 + H \rightarrow H + H_2$ reaction as a function of the collision energy. The black solid curve—calculations from [27] based on the surface with 8701 *ab initio* points from [29]. The dashed curves—calculations based on the GP PES obtained with 22 *ab initio* points (blue); 23 points (orange), 30 points (green) and 37 points (inset). The points 23–37 are drawn by the BO algorithm from the PES in [29]. The RMSE and maximum error of the results with 37 points are 0.009 and 0.028, respectively. Lower panel: the GP model of the PES for the H_3 reaction system constructed with 30 *ab initio* points. R_1 and R_2 are the distances between atoms 1 and 2 and atoms 2 and 3, respectively.

$$\sqrt{\{\mathcal{F}[\mathcal{G}(n+1,\boldsymbol{r})] - \mathcal{F}[\mathcal{G}(n)]\}^2}, \tag{17}$$

while restricting $\boldsymbol{r}$ to be within a physically reasonable range. This is justified by the observation that $\mathcal{G}(n \rightarrow \infty)$ must produce the best surface so the maximum improvement of the surface at each iteration is achieved when $\boldsymbol{r}$ corresponds to the maximum of equation (17).

In the subsequent sections, we illustrate these two approaches.

5. BO efficiency

We begin by illustrating the efficiency of BO for constructing the PES from feedback loops in equation (2). Because every cycle in equation (2) involves a quantum scattering calculation, the particular parameter of interest is the number of feedback loops necessary to produce a PES yielding accurate quantum dynamical results.

In this section, our goal is to obtain known reactive scattering probability for reactions (15) and (16), starting from a random guess of a PES and using as few cycles of equation (2) as possible. To do this, we train $\mathcal{F}[\mathcal{G}(n+1,\boldsymbol{r})]$ to represent the root mean square (RMS) deviation of the reaction probabilities from the known results. Here, BO amounts to minimizing the following function:

$$\mathcal{F} + \kappa\Delta\mathcal{F}. \tag{18}$$

The value of κ determines the efficiency of the optimization convergence. We set $\kappa = 0.005$. As described above, this value of κ is chosen by examining the first iteration of cycle (2) to ensure that the algorithm produces improved results with less than $15 \times \mathcal{N}$ dynamical calculations. After a local minimum of $\mathcal{F}$ is reached, the

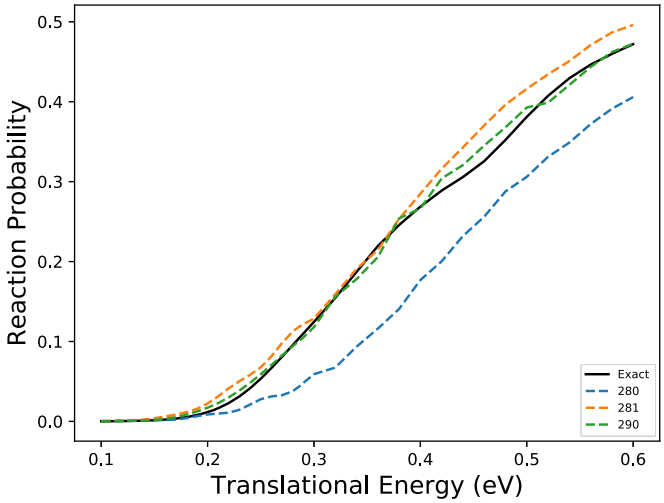


Figure 2. The reaction probability for the $OH + H_2 \rightarrow H + H_2O$ reaction as a function of the collision energy. The black solid curve —accurate calculations from [28] based on the surface constructed with ~17 000 *ab initio* points. The dashed curves—calculations based on the GP PES obtained with 280 *ab initio* points (blue); 281 points (orange) and 290 points (green). The points 281–290 are drawn by the BO algorithm from the PES in [28]. The RMSE and maximum error of the 290-point result are 0.0076 and 0.0195, respectively.

second term in equation (18) directs the minimization algorithm towards the parts of the configuration space where $\mathcal{F}$ is least accurate, thus sampling the entire space in search of the global minimum [32, 33]. With this procedure, the minimization of $\mathcal{F}$ for each value of n requires $15 \times \mathcal{N} = 45$ quantum dynamics calculations for H_3 and $15 \times \mathcal{N} = 90$ calculations for OH_3.

Figure 1(a) illustrates the performance of this algorithm in search of the best PES for reaction (5). As can be seen, the starting model $\mathcal{G}$ of the PES based on 22 randomly generated *ab initio* points produces highly inaccurate results, but the BO scheme converges to the correct PES after only 8 iterations ($8 \times 15 \times 3 = 360$ scattering calculations), yielding accurate results for the reaction probabilities (green dashed line) with the GP model $\mathcal{G}$ of the PES trained by 30 *ab initio* points. Figure 1(b) shows the model $\mathcal{G}$ of the PES obtained with $n = 30$ *ab initio* points, illustrating that equations (3) and (4) produce a physical surface.

Figure 2 illustrates the performance of this algorithm for the 6D reaction (6). As the dimensionality of the configuration space increases, so does n in equation (3) required to represent accurately the PES. Nevertheless, accurate results for the reaction probabilities (green dashed line) are obtained with $n = 290$ *ab initio* points, much smaller than the set of ~17 000 points used in previous work [28] to construct the PES with a NN fit. This result is obtained after 10 iterations (600 scattering calculations). The rms error of the reaction probabilities thus obtained is 0.0076. Note that, as any supervised learning technique, this algorithm is guaranteed to become more accurate when trained by more *ab initio* points.

The efficiency of BO is perhaps best illustrated by the dependence of the RMSEs of the reaction probabilities as functions of the number of BO iterations, presented in figure 3. Both panels of figure 3 show that BO quickly decreases the RMSE of the reaction probabilities. In the upper panel showing the results for reaction (5), the initial surface is constructed using 22 *ab initio* points chosen at random and BO starts with $n = 23$. In the lower panel showing the results for reaction (6), the initial surface is constructed using 280 *ab initio* points chosen at random in the 6D space and BO (results represented by circles) starts with $n = 281$. To illustrate quantitatively the accuracy gain enabled by BO, we also show in the lower panel of figure 3 the RMSE of the reaction probabilities computed with the PES obtained based on the corresponding number of points chosen at random and distributed in the configuration space using the Latin hypercube sampling scheme to avoid clustering. The difference in the computation difficulty between the results shown by squares and circles in the lower panel of figure 3 is 14 scattering calculations, i.e. the results shown by squares require one scattering calculations, whereas the results shown by circles require 15 scattering calculations. The key result illustrated by figure 3 is that the convergence of BO is fast and monotonous, reducing the error of the scattering calculations dramatically after only two to three iterations for both reactive systems. This underscores the critical importance of BO for the approach proposed here and also underlines the huge potential of BO for other applications involving the inverse scattering problem.

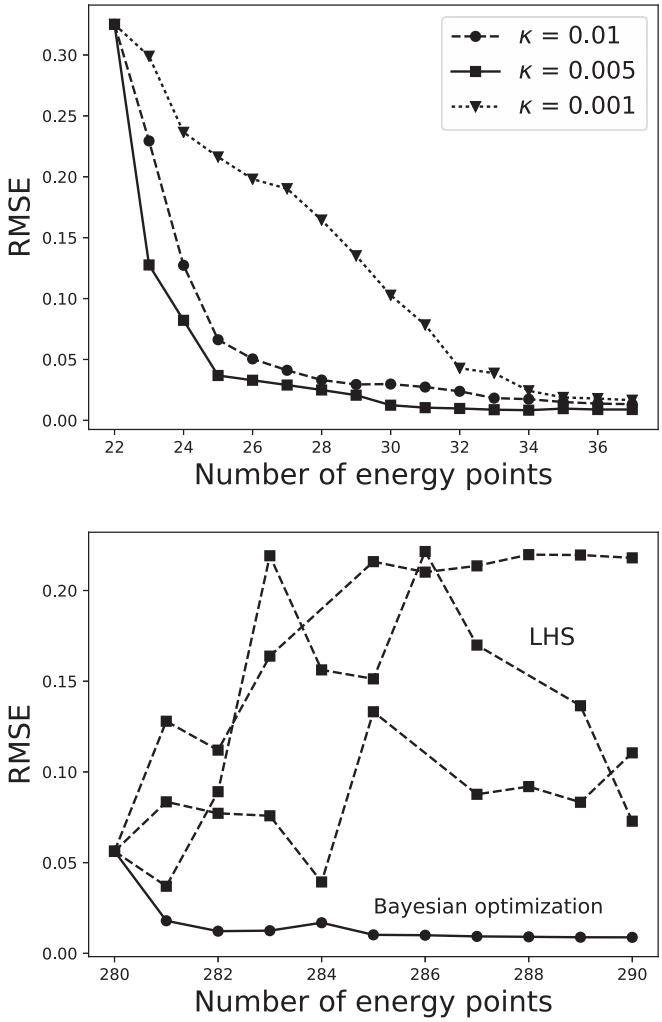


Figure 3. Convergence of the RMSE of the reaction probabilities. Upper panel: the $H_2 + H \rightarrow H + H_2$ reaction. Bayesian optimization starts at 22 points. The results shown by circles ($\kappa = 0.01$), squares ($\kappa = 0.005$) and triangles ($\kappa = 0.001$) illustrate the effect of the value of κ on convergence of Bayesian optimization. Lower panel: the $OH + H_2 \rightarrow H + H_2O$ reaction. Bayesian optimization of the results shown by circles starts at 280 points. The squares show the results of the scattering calculations with the PES obtained based on the corresponding number of points chosen at random and distributed in the configuration space using the Latin hypercube sampling scheme to avoid clustering.

6. Obtaining PES without known dynamical results

If the quantum dynamics results are not known, the present method can be applied to construct accurate PES with a small number of *ab initio* energy points. Since GP regression is a supervised learning algorithm, GP models of PES become necessarily more accurate when trained by more *ab initio* points (i.e. any interpolation method becomes infinitely accurate in the limit of infinite number of points to be interpolated). The method proposed here aims to reach this limit with as few steps as possible. To achieve this, it is necessary to restrict the configuration space of the molecule to a reasonable volume and maximize the difference between of the results of subsequent iterations in equation (2).

To illustrate the validity of this assumption, we show in figure 4 a series of computations as functions of n, showing the convergence of the iterative calculations to the accurate results (black solid curve). Figure 4 shows that the optimization loop converges to the accurate PES after 48 iterations. We emphasize that the accurate results (black curve) were not used in any way in this calculation.

7. Reproducing arbitrary scattering observables

Here, we extend the previous sections to construct a PES that, when used in quantum scattering calculations, reproduces an arbitrary set of observables. We first modify the exact scattering results of figure 1 by shifting along the energy axis and randomly modulating the black curve. This produces an arbitrary energy dependence

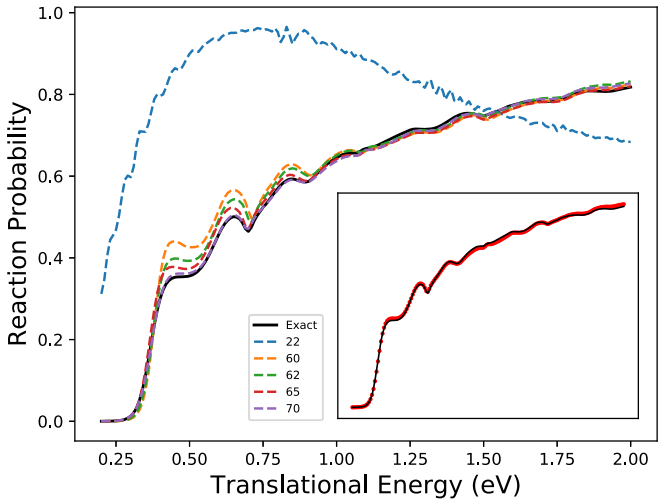


Figure 4. The reaction probabilities for the $H_2 + H \rightarrow H + H_2$ reaction as functions of the collision energy. The black solid curve—accurate calculations from [29]. The dashed curves—the results of iterative calculations maximizing the difference between the reaction probabilities in successive iterations. The black curve is not used for these calculations. The inset shows the agreement between the reaction probabilities (red symbols) based on the GP approach after 48 iterations (total of 70 *ab initio* points) and the exact results.

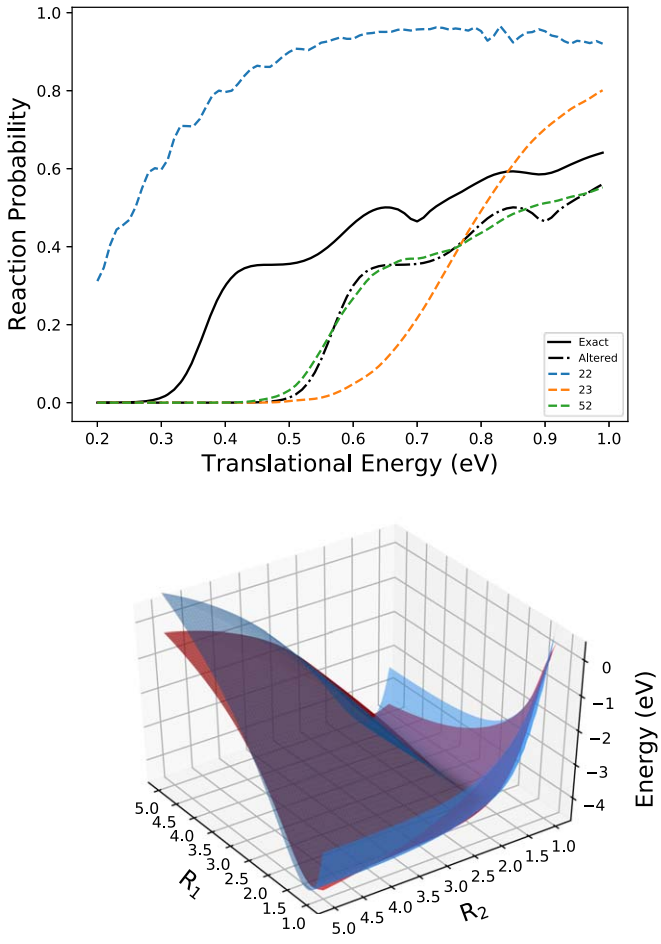


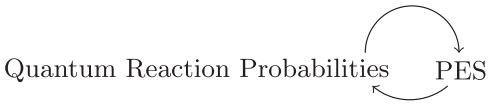
Figure 5. Upper panel: the reaction probabilities for the modified $H_2 + H \rightarrow H + H_2$ reaction as functions of the collision energy. The black dotted–dashed curve is obtained by a modification of the previous results (black solid curve) involving a translation along the energy axis. The ML models are trained to obtain the PES that would describe the new reaction probabilities. The green dashed curve is a results of such training after 30 iterations, resulting in a surface constructed with 52 *ab initio* points. Lower panel: comparison of the original PES (blue) with the new PES (red) found by the BO algorithm. The new PES yields the reaction probabilities described by the green dashed curve in the upper panel. The RMSE and maximum error of the results shown by the green dashed curve are 0.016 and 0.0674, respecively. R_1 and R_2 are defined as in figure 1.

of the reaction probabilities shown by the dotted–dashed curve in figure 5. The goal is to construct a PES that reproduces these arbitrarily chosen reaction probabilities. Note that the dotted–dashed curve extends the interval of energies, where the reaction probability is zero, which means that the PES for this reaction must have a higher reaction barrier and cannot be reproduced with the original PES for H_3. Since this approach builds the PES based on the observables and is designed to yield the PES reproducing the observables, it is equivalent to solving the inverse scattering problem.

We assume that a small ensemble of E_i is known from some (not necessarily accurate) quantum chemistry calculation. As before, this ensemble serves as a starting point for the model $\mathcal{G}$ of the PES. However, in order to allow for the improvement of the PES, we now allow the energy of each point of the PES to be a variable $\varepsilon(\boldsymbol{r})$. This variable is a function of $\boldsymbol{r}$ and allows the model $\mathcal{G}$ to sample from an interval of energies centered about the points of the initial rough PES. The results of the scattering calculations are now used to train the model $\mathcal{F}[\mathcal{G}(n+1, \varepsilon(\boldsymbol{r}), \boldsymbol{r})]$, effectively increasing dimensionality of the variable space to $\mathcal{N}+1$. Figure 5 shows that this algorithm converges to the arbitrarily modified reaction probabilities after 32 iterations, producing a PES depicted in the lower panel.

8. Conclusion

This work shows that it is possible to implement the feedback loop



to build the global PES for 3D and 6D reaction systems. The unique feature of this method is that the PES is built from a small number of *ab initio* points. At each iteration, an *ab initio* point is added to the most relevant part of the configuration space and the global PES is automatically morphed to become more accurate. By construction, this approach produces PES yielding an accurate description of known reaction observables and the PES offers unique information on which parts of the interaction are most important for the outcome of the reaction process.

By construction, this method illustrates the lowest number of potential energy points (i.e. the minimum information) required for the non-parametric cosntruction of global PES for quantum reactive scattering calculations. We showed that accurate quantum reactive scattering results can be obtained with 30 *ab initio* points for the 3D $H + H_2 \rightarrow H_2 + H$ reaction and 290 points for the 6D $OH + H_2 \rightarrow H + H_2O$. For practical applications, the locations of these few points can be determined as described above with a low-level *ab initio* method and an approximate dynamical approach. Once the positions of the points are known, this small number of *ab initio* points can be calculated with extremely high precision.

We emphasize that the method can also be used to construct accurate PES even if observables are not known. We showed that an algorithm based on *maximizing the difference* of computed observables in successive iterations quickly converges to an accurate PES without the *a priori* knowledge of the full surface or dynamical results. The convergence to the correct limit is guaranteed by the observation that the ML model adopted here produces the correct surface in the limit of a large number of *ab initio* points.

The efficiency of the approach proposed here is limited by the difficulty of the dynamical calculations. However, one can use approximate dynamical methods for step (iii) in equation (2). This will result in PES, which yields an accurate description of the observables if used with this specific dynamical method. The error of this PES will be designed to compensate for the error of the dynamical calculations. While it remains to be seen if such an approach has predictive power outside the range of the observables used to construct the PES, we believe, our work opens the possibility of applying approximate dynamical methods such as the coupled states approach [34–38] for quantitative predictions of molecular scattering observables.

Finally, we note that the approach presented here is general and can be used to construct the microscopic interaction potentials for any process with a

$$\text{Microscopic Interaction Potential} \Rightarrow \text{Observable}$$

dependence. Within the framework of the method proposed here, one should represent the interaction potential by a GP model $\mathcal{G}$ and the observable by another GP model $\mathcal{F}[\mathcal{G}]$. The GP $\mathcal{F}[\mathcal{G}]$ can then be used as the emulator function for BO to find the optimal model $\mathcal{G}$, leading to the most accurate description of the observable. We believe this strategy can be applied to a wide range of problems in molecular physics.

Acknowledgments

This work is supported by NSERC of Canada. RVK also acknowledges support from the Chinese Academy of Sciences in the form of a visiting fellowship. DHZ acknowledges support from the National Natural Science Foundation of China (grants 21433009, 21590804, 21688102), and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB17000000).

References

[1] Krems R V 2008 *Phys. Chem. Chem. Phys.* **10** 4079
[2] Ernesti A and Hutson J M 1997 *J. Chem. Phys.* **106** 6288–301
[3] Meuwly M and Hutson J M 1999 *J. Chem. Phys.* **110** 8338–47
[4] Howson J M M and Hutson J M 2001 *J. Chem. Phys.* **115** 5059–65
[5] Cho Y-S and Le Roy R J 2016 *J. Chem. Phys.* **144** 024311
[6] Li Z, Singh S, Tscherbul T V and Madison K W 2008 *Phys. Rev.* A **78** 022710
[7] Murrell J N, Carter S, Farantos S C, Huxley P and Varandas A J C 1984 *Molecular Potential Energy Functions* (Chichester: Wiley)
[8] Hollebeek T, Ho T-S and Rabitz H 1999 *Annu. Rev. Phys. Chem.* **50** 537
[9] Braams B J and Bowman J M 2009 *Int. Rev. Phys. Chem.* **28** 577
[10] Collins M A 2002 *Theory Chem. Acc.* **108** 313
[11] Handley C M and Popelier P L A 2010 *J. Phys. Chem.* A **114** 3371
[12] Handley C M, Hawe G I, Kellab D B and Popelier P L A 2009 *Phys. Chem. Chem. Phys.* **11** 6365
[13] Bartók A P, Payne M C, Kondor R and Csányi G 2010 *Phys. Rev. Lett.* **104** 136403
[14] Bartok A P and Csanyi G 2015 *Int. J. Quant. Chem.* **115** 1051
[15] Caccin M, Li Z, Kermode J R and De Vita A 2015 *Int. J. Quant. Chem.* **115** 1129
[16] Popelier P L A 2015 *Int. J. Quant. Chem.* **115** 1005
[17] Cui J and Krems R V 2016 *J. Phys. B: At. Mol. Opt. Phys.* **49** 224007
[18] Kamath A, Vargas-Hernández R A, Krems R V, Carrington T Jr and Manzhos S 2018 *J. Chem. Phys.* **148** 241702
[19] Kolb B, Marshall P, Zhao B, Jiang B and Guo H 2017 *J. Phys. Chem.* A **121** 2552
[20] Qu C, Yu Q, Van Hoozen B L, Bowman J M and Vargas-Hernndez R A 2018 *J. Chem. Theory Comput.* **14** 3381
[21] Rasmussen C E and Williams C K I 2006 *Gaussian Processes for Machine Learning* (Cambridge, MA: MIT Press)
[22] Cui J and Krems R V 2015 *Phys. Rev. Lett.* **115** 073202
[23] Qiu M H *et al* 2006 *Science* **311** 1440
[24] Che L *et al* 2007 *Science* **317** 1061
[25] Dong W, Xiao C, Wang T, Dai D, Yang X and Zhang D H 2010 *Science* **327** 1501
[26] Wang T, Chen J, Yang T, Xiao C, Sun Z, Huang L, Dai D, Yang X and Zhang D H 2013 *Science* **342** 1499
[27] Su N Q, Chen J, Sun Z, Zhang D H and Xu X 2015 *J. Chem. Phys.* **142** 084107
[28] Chen J, Xu X, Xu X and Zhang D H 2013 *J. Chem. Phys.* **138** 154301
[29] Boothroyd A I, Keogh W J, Martin P G and Peterson M R 1996 *J. Chem. Phys.* **104** 7139
[30] Sutton R S and Barto A G 2016 *Reinforcement Learning, An Introduction* (Cambridge, MA: MIT Press)
[31] Settles B 2010 Active learning literature survey *Comput. Sci. Technol. Rep.* **1648**
[32] Snoek J, Larochelle H and Adams R P 2012 *Adv. Neural Inf. Process. Syst.* **25** 2951
[33] Shahriari B, Swersky K, Wang Z, Adams R P and de Freitas N 2016 *Proc. IEEE* **104** 148
[34] Pack R T 1974 *J. Chem. Phys.* **60** 633
[35] McLenithan K and Secrest D 1984 *J. Chem. Phys.* **80** 2480
[36] Khare V, Kouri D J and Pack R T 1978 *J. Chem. Phys.* **69** 4419
[37] McGuire P and Kouri D J 1974 *J. Chem. Phys.* **60** 2488
[38] McGuire P 1975 *J. Chem. Phys.* **62** 525