

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.




<http://hdl.handle.net/2066/201526>

Please be advised that this information was generated on 2020-09-10 and may be subject to change.

Six-dimensional potential energy surface for NaK–NaK collisions: Gaussian process representation with correct asymptotic form

Cite as: J. Chem. Phys. **150**, 064106 (2019); <https://doi.org/10.1063/1.5082740>

Submitted: 23 November 2018 . Accepted: 25 January 2019 . Published Online: 13 February 2019

Arthur Christianen , Tijs Karman , Rodrigo A. Vargas-Hernández, Gerrit C. Groenenboom , and Roman V. Krems



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Operators in quantum machine learning: Response properties in chemical space](#)

The Journal of Chemical Physics **150**, 064105 (2019); <https://doi.org/10.1063/1.5053562>

[Vibronically and spin-orbit coupled diabatic potentials for \$X\(P\) + CH_4 \rightarrow HX + CH_3\$ reactions:](#)

[General theory and application for \$X\(P\) = F\(^2P\)\$](#)

The Journal of Chemical Physics **150**, 064102 (2019); <https://doi.org/10.1063/1.5063907>

[Invariance of experimental observables with respect to coarse-graining in standard and many-body dissipative particle dynamics](#)

The Journal of Chemical Physics **150**, 064101 (2019); <https://doi.org/10.1063/1.5046851>

Where in the **world** is AIP Publishing?

Find out where we are exhibiting next



Six-dimensional potential energy surface for NaK–NaK collisions: Gaussian process representation with correct asymptotic form

Cite as: J. Chem. Phys. 150, 064106 (2019); doi: 10.1063/1.5082740

Submitted: 23 November 2018 • Accepted: 25 January 2019 •

Published Online: 13 February 2019



View Online



Export Citation



CrossMark

Arthur Christianen,¹  Tijs Karman,²  Rodrigo A. Vargas-Hernández,³ Gerrit C. Groenenboom,¹ 
and Roman V. Krems^{3,a)}

AFFILIATIONS

¹Theoretical Chemistry, Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

²ITAMP, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA

³Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

^{a)}Author to whom correspondence should be addressed: rkrems@chem.ubc.ca

ABSTRACT

Constructing accurate global potential energy surfaces (PESs) describing chemically reactive molecule-molecule collisions of alkali metal dimers presents a major challenge. To be suitable for quantum scattering calculations, such PESs must represent accurately three- and four-body interactions, describe conical intersections, and have a proper asymptotic form at the long range. Here, we demonstrate that such global potentials can be obtained by Gaussian Process (GP) regression merged with the analytic asymptotic expansions at the long range. We propose an efficient sampling technique, which allows us to construct an accurate global PES accounting for different chemical arrangements with <2500 *ab initio* calculations. We apply this method to (NaK)₂ and obtain the first global PES for a system of four alkali metal atoms. The resulting surface exhibits a complex landscape including a pair and a quartet of symmetrically equivalent local minima and a seam of conical intersections. The dissociation energy found from our *ab initio* calculations is 4534 cm⁻¹. This result is reproduced by the GP models with an error of less than 3%. The GP models of the PES allow us to analyze the features of the global PES, representative of general alkali metal four-atom interactions. Understanding these interactions is of key importance in the field of ultracold chemistry.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5082740>

I. INTRODUCTION

A major thrust of recent experimental research in molecular physics has been to create a gas of stable diatomic molecules at ultracold temperatures.^{1,2} Multiple experiments have demonstrated the formation of ultracold molecules by photoassociation of ultracold atoms.^{3–9} Here we consider the NaK molecule. The reaction $\text{NaK} + \text{NaK} \rightarrow \text{Na}_2 + \text{K}_2$ is endothermic,¹⁰ and NaK molecules are therefore believed to be chemically stable at ultracold temperatures. However, the experiments with ultracold NaK molecules report losses,⁸ which have been attributed to the formation of long-lived

collision complexes.^{11,12} In order to understand the origin of these losses, it is necessary to perform quantum scattering calculations of cross sections for NaK–NaK collisions. Such calculations require an accurate global potential energy surface (PES) for the NaK–NaK collision complex.

Constructing this PES is a complex task. The endoergicity of the $\text{NaK} + \text{NaK} \rightarrow \text{Na}_2 + \text{K}_2$ reaction is only 120 cm⁻¹, whereas the dissociation energy of the NaK–NaK complex is calculated here to be 4534 cm⁻¹. This implies that a global PES suitable for scattering calculations must account for the configuration space corresponding to the exchange of atoms. At the same time, the PES must describe accurately the three- and

four-body interactions and have a proper asymptotic form. While high-level accurate *ab initio* calculations for $(\text{NaK})_2$ are feasible, it is a major challenge to construct an analytical fit of the PES describing fully and accurately the complexity of NaK–NaK interactions in a wide range of energies (from the ultracold collision asymptote, to the reactive complex, to all reaction channels) required for ultracold scattering calculations.

An alternative to an analytical fit is a non-parametric machine-learning (ML) model, such as Gaussian Process (GP) regression.¹³ This method provides an *interpolation* of *ab initio* points that does not rely on any functional form to represent the PES and is, therefore, completely general. The PES is obtained as a mean of a multi-variate Gaussian distribution. This generality makes GP regression a very useful and powerful method for obtaining the PES. For example, the same code without any additional programming effort can be applied to construct a PES for systems with different atoms (e.g., RbCs vs NaK) and, even, with different dimensions (e.g., NaK–NaK vs Na–NaK). As shown recently,^{14–18} GP regression can produce accurate PESs for polyatomic systems with a small number of *ab initio* points. However, as any ML method, GP regression is designed for interpolation and cannot extrapolate outside the range of available *ab initio* energy points. The conventional implementation of GP regression is, thus, not suitable for describing the asymptotic behaviour of the PES that is critically important for ultracold scattering calculations.

The present work pursues three goals: (i) extend the previous work on GP regression to design a general method for constructing a multi-dimensional reactive PES with correct asymptotic behaviour suitable for ultracold scattering calculations (Secs. II, III, and IV B); (ii) obtain an accurate PES for the NaK–NaK interactions; and (iii) examine the general features of this PES (Sec. IV A), which we expect to be characteristic of PESs for other alkali metal tetramers in the ground electronic state.

II. AB INITIO CALCULATIONS

The coordinate system most suitable for a non-reactive diatom-diatom system is the set of Jacobi coordinates depicted in Fig. 1. At the long range in the NaK–NaK asymptotic arrangement, we use these Jacobi coordinates and an associated Legendre expansion to describe the dependence of the PES on the polar angles θ_1 , θ_2 and the dihedral angle ϕ .

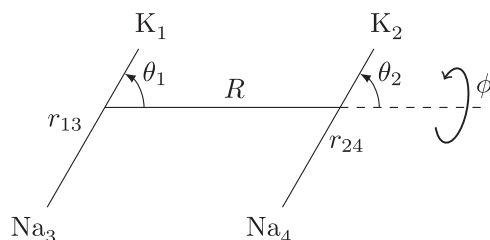


FIG. 1. The Jacobi coordinates of the complex in one of the NaK–NaK arrangements.

This expansion diverges at the short range. Furthermore, the Jacobi coordinates do not describe the different chemical arrangements of the four-atom complex equivalently, making it difficult to fit the reactive part of the PES. Therefore, at the short range, we use the inverse atomic distance coordinates.

In the region where the Jacobi coordinate $R < 25 a_0$, the *ab initio* points were calculated using a configuration interaction (CI) method. At $R \geq 15 a_0$, the *ab initio* points were calculated using coupled-cluster theory. Note that there is an overlap between both regions, to ensure a smooth transition between the two areas. The computational approaches for both regions are summarized below. For all calculations, we employ effective core potentials (ECPs) and core polarization potentials (CPPs) for Na and K such that only the four valence electrons are treated explicitly. We use MOLPRO¹⁹ for all the *ab initio* calculations.

We use the effective core potentials (ECPs) developed by Nicklass *et al.*,²⁰ to represent all but a single valence electron for each sodium and potassium atom. The core polarizability (and therefore also most of the core-valence correlation²¹) is accounted for using the CPPs developed by Fuentealba *et al.*²² and implemented in MOLPRO.¹⁹ The CPP parameters are the dipole polarizability $\alpha = 0.9947 (5.354)a_0^3$ and the exponential cutoff parameter $\delta = 0.62 (0.29)a_0$, in atomic units, for sodium (potassium).²² First, we roughly optimized the one-electron basis sets as follows. We use uncontracted Gaussian orbitals throughout. The *s*, *p*, and *d* orbitals are represented by identical sets of five even-tempered orbitals. Two additional *f* atomic orbitals are included. The exponents were very roughly optimized by variationally lowering the energy of the NaK–NaK dimer at a head-to-tail collinear geometry. We chose not to optimize the basis sets for the one-electron atoms because this corresponds to a special case with no electron correlation and effectively optimizes only the included valence orbitals. This procedure leads to Gaussian exponents (1.0, 0.3162, 0.1, 0.0316, 0.01) for the *s*, *p*, *d* orbitals and (0.08, 0.008) for the *f* orbitals.

When applied to the isolated Na or K atom, the atomic energy levels can be obtained from Hartree-Fock calculations for the effective one-electron atoms. We perform such calculations for both 2S and 2P atomic states, which yields the valence $S \rightarrow P$ transitions at 16 826.85 (13 015.17) cm^{-1} for Na (K). These agree to an accuracy of 0.8 (0.07)% with the experimental²³ transition wave number of 16 967.63 (13 023.66) cm^{-1} . Note that the error on the $S \rightarrow P$ transition of the Na-atom is a factor 10 higher than the error for the K-atom. After completing the PES, we performed additional tests of the one electron basis. We independently optimized the Na and K basis sets on the minima of the Na_2 and K_2 diatoms by scaling the exponents. This resulted in roughly two times bigger exponents for Na than used now, whereas the exponents for K were already optimal. Choosing this basis set for Na would reduce the error of the Na $S \rightarrow P$ transition to 0.04% and lower the minimum energy of the PES by 30 cm^{-1} . Since the effect of this basis set change is smaller than the typical GP fitting error such as shown in Tables III and V, we did not recompute the PES.

TABLE I. Spectroscopic parameters of NaK.

State	Parameter	This work	Experiment	References	Error (%)
$X^1\Sigma^+$	D_e/cm^{-1}	5319.89	5273.62	24	0.9
	R_e/a_0	6.59	6.61	24	0.3
	$\hbar\omega/\text{cm}^{-1}$	122.79	124.01	24	1.0
$a^3\Sigma^+$	D_e/cm^{-1}	209.81	207.82	24	1.0
	R_e/a_0	10.26	10.29	24	0.3
	$\hbar\omega/\text{cm}^{-1}$	22.715	23.01	24	1.3
$b^3\Pi$	D_e/cm^{-1a}	6869.68	6735.10	25	2.0
	R_e/a_0	6.59	6.62	25	0.5
	$\hbar\omega/\text{cm}^{-1}$	120.19	119.70	25	0.4

^aRelative to the Na(3s) + K(4p) asymptote.

Next, we calculated the potential energy curves for the two-electron NaK molecule at the full (single and double) CI level. The spectroscopic parameters obtained from these calculations are compared to experimental results in Table I. Most of the spectroscopic parameters are found to be accurate to within one percent.

For the four-valence electron NaK–NaK dimer, we find that the computational cost of the full CI approach increases too rapidly to obtain convergence with respect to the one-electron basis set. Therefore, we use truncated CI and coupled-cluster approaches as outlined below.

In the short range, multi-reference CI calculations were performed as follows. The ECPs and CPPs discussed above are employed throughout. A complete-active space self-consistent field (SCF) calculation was performed for the four valence electrons. Subsequently, we perform an internally contracted multireference CI step, where single and double excitations from this active space are taken into account. We applied the Pople size-consistency correction.²⁶ We note that we find that the Pople correction is much more accurate than the more commonly used Davidson correction²⁷ for this system, which contains only four explicit electrons. We also include the Boys and Bernardi counterpoise correction²⁸ to reduce the basis set superposition error by subtracting the energies of the four separate atoms calculated in the dimer basis set. These were calculated at a Hartree-Fock level, which gives the same results as the full CI for these effectively one-electron atoms.

In the long-range, we performed coupled-cluster calculations including single and double excitations and perturbative triples [CCSD(T)], based on Hartree-Fock reference determinants. This again treats only the four valence electrons and effectively accounts for the core electrons using ECPs and CPPs. This approach gives a more accurate account of the long-range interaction but is not applied in the reactive short-range region where the single-determinant description is not generally applicable.

III. FITTING THE GLOBAL PES

The general strategy used in this work aims to merge a non-parametric machine-learning model of the PES at the short range with a proper analytical form of the PES at the long

range. In order to achieve this, one must overcome several issues, discussed below.

Note that throughout the paper we use the energy of two isolated NaK diatoms at their equilibrium bond lengths as the zero of energy.

A. GP regression

Gaussian process (GP) regression is a kernel-based probabilistic non-parametric supervised ML algorithm.¹³ The application of GP regression to constructing a PES for polyatomic molecules has been described in detail in the previous work of one of the authors.¹⁴ The GP model is trained by an ensemble of *ab initio* points distributed in the relevant configuration space and provides a prediction of the energy as a normal distribution characterized by a mean $\mu(\cdot)$ and a standard deviation $\sigma(\cdot)$ in the following form:

$$\mu(\mathbf{x}_*) = K(\mathbf{x}_*, \mathbf{x})^\top [K(\mathbf{x}, \mathbf{x}) + \sigma_n^2 I]^{-1} \mathbf{y}, \quad (1)$$

$$\sigma(\mathbf{x}_*) = K(\mathbf{x}_*, \mathbf{x}_*) - K(\mathbf{x}_*, \mathbf{x})^\top [K(\mathbf{x}, \mathbf{x}) + \sigma_n^2 I]^{-1} K(\mathbf{x}, \mathbf{x}_*). \quad (2)$$

Here, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$ is a vector of n points in the six-dimensional configuration space of the NaK–NaK complex, where the GP model is trained, \mathbf{y} is a vector of *ab initio* points at the positions specified by \mathbf{x}_i , \mathbf{x}_* is a point in the configuration space where the prediction \mathbf{y}_* is to be made, $K(\mathbf{x}, \mathbf{x})$ is the $n \times n$ square matrix with the elements $K_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$ representing the covariances between $\mathbf{y}(\mathbf{x}_i)$ and $\mathbf{y}(\mathbf{x}_j)$, and $k(\cdot, \cdot)$ is a kernel function. The purpose of training a GP model is to find the parameters of the kernel where the log *marginal likelihood*,

$$\log p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = -\frac{1}{2} \mathbf{y}^\top K^{-1} \mathbf{y} - \frac{1}{2} \log |K| - \frac{n}{2} \log(2\pi), \quad (3)$$

is maximum. Here, $\boldsymbol{\theta}$ denotes collectively the parameters of the analytical function for $k(\cdot, \cdot)$ (see below) and $|K|$ is the determinant of the matrix K .

In the present work, we assume the following form for the kernel function:

$$k(\mathbf{x}_i, \mathbf{x}_j) = cD(\mathbf{x}_i, \mathbf{x}_j, \sigma)M(\mathbf{x}_i, \mathbf{x}_j, l_1, l_2, l_3, l_4, l_5, l_6), \quad (4)$$

where D is the dot product kernel¹³

$$D(\mathbf{x}_i, \mathbf{x}_j, \sigma) = \sigma^2 + \mathbf{x}_i \cdot \mathbf{x}_j, \quad (5)$$

and M is an anisotropic Matérn kernel¹³ with $\nu = 2.5$,

$$M(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k=1}^6 \left(1 + \sqrt{5} l_k^{-1} |\mathbf{x}_{i,k} - \mathbf{x}_{j,k}| + \frac{5}{3} l_k^{-2} |\mathbf{x}_{i,k} - \mathbf{x}_{j,k}|^2 \right) \times \exp\left(-\sqrt{5} l_k^{-2} |\mathbf{x}_{i,k} - \mathbf{x}_{j,k}|^2\right). \quad (6)$$

The coefficient c , the dot product bias term σ , and the characteristic length scales l_1 – l_6 of the Matérn function are the parameters $\boldsymbol{\theta}$ in Eq. (3).

The parameters of the kernels and the training points that were used to make the GP fits of the NaK–NaK PES are given in the [supplementary material](#).

The accuracy of the GP interpolation increases monotonically with the number n of training points.¹⁴ However, the numerical evaluation of the PES involves a single inversion of a square $n \times n$ matrix and can subsequently be reduced to a vector-vector product for each evaluated point, scaling as $\mathcal{O}(n)$. For practical applications, it is therefore important to reduce n by placing the training points optimally in the configuration space. At the same time, the GP model has no prior information about the function to be fitted, such as the analytical form of the long range and the symmetry of the PES. To incorporate those properties correctly, the GP model needs to be symmetrized (Sec. III C) and it needs to be merged with the analytical long-range part of the PES (Sec. III D).

To optimize the quality of the fit, both the type of kernel and the coordinate representation (here inverse atomic distances) are important. In previous work, only a single Matérn kernel (with $\nu = 2.5$) was used.¹⁴ It is generally expected that more complex kernels can provide better results. Duvenaud *et al.*^{29,30} suggested that the complexity of kernels can be built up by combining simple kernels (such as the Matérn kernel or a simple Gaussian kernel) in the form of products and sums using the Bayesian information criterion computed for the resulting model as a kernel selection criterion. This procedure was recently shown to produce kernels, allowing for physical extrapolation of quantum properties with Gaussian processes.³¹ Here, we build up the complexity of kernels using the same algorithm but with the testing error instead of the Bayesian information criterion to guide the kernel selection. The final kernels thus obtained were dependent on the training set, as is expected and was also discussed in Ref. 31. However, the resulting kernels always contained a product of a Matérn kernel and a dot product kernel and the other components of the obtained kernels only marginally reduced the fitting error. We therefore chose to use this product of a Matérn kernel ($\nu = 2.5$) and a dot product kernel as our kernel for constructing the PES, as shown in Eq. (4). This kernel produces functions which are twice differentiable. The maximization of the log-likelihood function was performed with the Python(2.7) package scikit-learn 0.18.1.³²

B. Configuration space sampling

To place the training points for the GP model, the following procedure was used. A grid in Jacobi coordinates is constructed using Latin hypercube sampling (LHS).³³ For the NaK–NaK arrangement, this is performed using a grid in the monomer bond lengths r_{13} and r_{24} from 4.6 to 9.4 a_0 and with $r_{24} \geq r_{13}$. The labeling of the atoms is defined in Fig. 1, and r_{ij} is the distance between atoms i and j . The range of R-values used is 4 to 25 a_0 . The angles θ_1, θ_2 range from 0 to π . The angle ϕ ranges from 0 to 2π , but because of the symmetry only ϕ values up to π need to be included. For the Na₂–K₂ arrangement, the grid has $4.8 < r_{12} < 10.3 a_0$, $3.7 < r_{34} < 8.1 a_0$, $4 < R < 25 a_0$, $0 < \theta_1, \theta_2 < \frac{\pi}{2}$, and $0 < \phi < \pi$. The grid ranges are chosen such that almost the entire relevant parameter space for cold collisions is included. However, the grid ranges do

not depend on each other, meaning that a large part of the probed parameter space is not accessible. Since the number of points included in the GP training set affects the computational cost of the PES evaluation, only the *ab initio* points relevant for the dynamics should be included in the training set. It is thus desirable to develop an approach to determine whether a point should be included in the training set or not.

To select the training points, we use a computationally cheap PES constructed with the Diatomics-in-Molecules (DIM) model.³⁴ The diatom potentials used for this model were constructed as described in Sec. II. The DIM model is evaluated at each grid point. If the calculated energy of the grid point is below the selected cutoff energy, an *ab initio* point is calculated and included in the training set of the GP model. The number of points of the LHS grid was modified to obtain the desired number of remaining training points.

The cutoff energy is dependent on the bond lengths of the diatoms. Furthermore, it is chosen such that no points are wasted to fit the dissociation of the diatoms, but that a repulsive wall is still included when the diatoms approach each other. To choose the cutoff energy in the different regions in parameter space, we devised the following two rules:

1. If both monomer bond lengths in either the NaK–NaK or the Na₂K₂ arrangements are close to their equilibrium bond length ($\pm 0.5a_0$), the cutoff energy is twice the dissociation energy of the NaK diatoms.
2. If one of the bond lengths differs more than $0.5a_0$ from the equilibrium bond length, the cutoff energy is 624 cm^{-1} .

When all *ab initio* points on this grid are calculated, the coordinates are transformed to inverse atomic distances. These inverse atomic distances are used to produce a GP model of the surface. We find that the length scales $l_1 - l_6$ of the trained kernels [Eq. (4)] are circa $0.3 a_0^{-1}$ [see the [supplementary material](#)]. The inverse atomic distances vary from $\sim 0 a_0^{-1}$ to approximately $0.25 a_0^{-1}$. The length scales of the GP fits therefore match the length scale of the PES.

C. Symmetrization of the fit

There are four ways in which the nuclei of the NaK–NaK complex can be permuted which yield symmetrically equivalent arrangements. We can define permutation operators \hat{P}_{ij} , which interchange the nuclei i and j . The permutations \hat{I} (the identity), \hat{P}_{12} , \hat{P}_{34} , and $\hat{P}_{12}\hat{P}_{34}$ should leave the energy invariant. In this work, we train the GP models with points that are predominantly in one quadrant of the parameter space. Adding all symmetrically equivalent points to the Gaussian process model would unnecessarily increase the set of training points by a factor of four and therefore make the evaluation of the PES more computationally expensive. To avoid this, we developed the following symmetrization procedure. The fitted values corresponding to the four permutations are added with a weight determining whether the permuted set of coordinates corresponds to a geometry inside or outside the training point region.

In the Na₂-K₂ arrangement, the training point region has the corresponding Jacobi coordinate values $\theta_1, \theta_2 \in [0, \frac{\pi}{2}]$. In the NaK-NaK arrangement, the training point region has $r_{12} < r_{34}$ and most of the training point region also has $r_{13} + r_{24} < r_{23} + r_{14}$ because 1-3 and 2-4 are the diatoms in the chosen NaK-NaK arrangement. The symmetrization procedure is based on pairing different contributions using a switching function,

$$F_{\text{merge}}[u; m, w, F_1(\vec{x}), F_2(\vec{x})] = y(u, c, w)F_1(\vec{x}) + [1 - y(u, c, w)]F_2(\vec{x}), \quad (7)$$

$$y(u; c, w) = \begin{cases} 0, & \text{if } u \leq c - w \\ \frac{1}{2} + \frac{9}{16} \sin \frac{\pi(u-c)}{2w} + \frac{1}{16} \sin \frac{3\pi(u-c)}{2w}, & \text{if } c - w < u < c + w \\ 1, & \text{if } u \geq c + w. \end{cases} \quad (8)$$

Here $F_1(\vec{x})$ and $F_2(\vec{x})$ are the functions between which the switch is made, u is the parameter parameterizing the switch, c is the value of u around which the switch is centered, and w is the halfwidth of the switching interval. The sigmoid function y is constructed to switch in the finite interval $c - w < u < c + w$ and to be twice differentiable. For F_1 and F_2 , we use the GP predictions evaluated at symmetrically equivalent geometries.

In the Na₂-K₂ region, the symmetrization scheme is as follows:

$$V_{12}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\theta_1; \frac{\pi}{2}, 0.3, V^{\text{GP}}(\vec{x}), V^{\text{GP}}(\hat{P}_{12}\vec{x})], \quad (9)$$

$$V_{34}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\theta_1; \frac{\pi}{2}, 0.3, V^{\text{GP}}(\hat{P}_{34}\vec{x}), V^{\text{GP}}(\hat{P}_{12}\hat{P}_{34}\vec{x})], \quad (10)$$

$$V_{\text{Na}_2\text{K}_2}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\theta_2; \frac{\pi}{2}, 0.3, V_{12}^{\text{GP}}(\vec{x}), V_{34}^{\text{GP}}(\vec{x})]. \quad (11)$$

Then, in the NaK-NaK region,

$$V_{14}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\frac{r_{13}}{r_{24} + r_{13}}; \frac{1}{2}, \frac{1}{16}, V^{\text{GP}}(\vec{x}), V^{\text{GP}}(\hat{P}_{12}\hat{P}_{34}\vec{x})], \quad (12)$$

$$V_{23}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\frac{r_{23}}{r_{14} + r_{23}}; \frac{1}{2}, \frac{1}{16}, V^{\text{GP}}(\hat{P}_{12}\vec{x}), V^{\text{GP}}(\hat{P}_{34}\vec{x})], \quad (13)$$

$$V_{\text{NaKNaK}}^{\text{GP}}(\vec{x}) = F_{\text{merge}}[\frac{r_{13} + r_{24}}{r_{23} + r_{14} + r_{13} + r_{24}}; \frac{1}{2}, \frac{1}{16}, V_{14}^{\text{GP}}(\vec{x}), V_{23}^{\text{GP}}(\vec{x})]. \quad (14)$$

Finally, we pair up the functions of the two arrangements,

$$V_{\text{sym}}^{\text{GP}}(\vec{x}) = F_{\text{merge}} \left[\frac{r_{12} + r_{34}}{2(r_{13} + r_{24})} + \frac{r_{12} + r_{34}}{2(r_{14} + r_{23})}; 1, \frac{1}{4}, V_{\text{Na}_2\text{K}_2}^{\text{GP}}(\vec{x}), V_{\text{NaKNaK}}^{\text{GP}}(\vec{x}) \right]. \quad (15)$$

To ensure a smooth symmetrization, we need to add some symmetrically equivalent points to the training set at the edges of the quadrant of parameter space used for training. There are two NaK-NaK arrangements, but only one of the arrangements is probed by the GP. To obtain a correctly shaped well and to ensure the potential wall is still included after symmetrization, points up to $R = 10a_0$ for the second

NaK-NaK arrangement are also included. This was performed by permuting the interatomic distances of the already calculated *ab initio* points by \hat{P}_{34} and adding these new points as extra training points. In the Na₂-K₂ arrangement, points with θ_1 and θ_2 between $\pi/2$ and $\pi/2 + 0.2$ were included.

D. Fitting the long range

The GP and CI methods do not provide a correct description of the long range, so a different method was used to construct the long range of the PES. Analytical expansion coefficients describing the long range are available in the literature.³⁵ However, to make a smooth transition with the short-range part, higher-order coefficients and the bond-length dependencies of the coefficients are also needed. To calculate these higher coefficients and bond-length dependencies, we calculated *ab initio* points using CCSD(T). Then we extracted the Legendre expansion coefficients and fitted the difference between these expansion coefficients with the literature coefficients using the 1D reciprocal power reproducing kernel Hilbert space method (RKHS).³⁶

The *ab initio* points for both arrangements were calculated at intermolecular distances of $R = 15, 16, 17, 18, 20, 22, 25,$ and $30 a_0$. Note that at distances of more than $30 a_0$, the expansion coefficients displayed unphysical behaviour. We have indications that this is caused by the use of the exponentially decaying CPP's. A 6-point Gauss-Legendre quadrature was used for the coordinates θ_1 and θ_2 , and a 6-point Gauss-Chebyshev quadrature was used for ϕ . The long range is represented analytically by the following expansion:

$$V_{\text{lr}}(r_{13}, r_{24}, R, \theta_1, \theta_2, \phi) = \sum_{l_1, l_2=0}^4 \sum_{m=0}^{\min(l_1, l_2)} W_{l_1, l_2, m}(r_1, r_2, R) \times P_{l_1}^m[\cos(\theta_1)] P_{l_2}^m[\cos(\theta_2)] \cos(m\phi). \quad (16)$$

Here P_l^m are associated Legendre polynomials.³⁷ To also treat the r_1 and r_2 dependencies, $W_{l_1, l_2, m}(r_1, r_2, R)$ can be approximated as $W_{l_1, m}^{(1)}(r_1, R) W_{l_2, m}^{(2)}(r_2, R)$. This is exact for all electrostatic and induction terms because they depend on a product of monomer properties and approximate for the dispersion terms. This approximation is invalid, though, for terms where $W_{l_1, m}^{(1)}(r_1, R)$ or $W_{l_2, m}^{(2)}(r_2, R)$ goes through zero as a function of R . For this PES, this only occurs for some terms with l_1 or $l_2 = 3$ or 4 . For these terms, the r_1 - and r_2 -dependence is neglected. Since these terms are small, this has only a small effect on the quality of the fit. The r_1 and r_2 dependence was approximated by second degree polynomials in r_1 and r_2 , i.e.,

$$W_{l_1, l_2, m}(r_1, r_2, R) = \sum_{k_1, k_2=0}^2 W_{l_1, m, k_1}^{(1)}(R) W_{l_2, m, k_2}^{(2)}(R) (r_1 - r_{1\text{eq}})^{k_1} (r_2 - r_{2\text{eq}})^{k_2}. \quad (17)$$

The dominant interaction coefficients were calculated using the multipole moments and dispersion coefficients

reported in Table II. For the electrostatic terms, the W -coefficients can be calculated from these multipole moments Q_l such as described by Byrd,³⁵

$$W_{l_1, l_2, m}(r_1, r_2, R) = (-1)^{l_1+m} (2 - \delta_{m,0}) \frac{(l_1 + l_2)!}{(l_1 + m)!(l_2 + m)!} \times Q_{l_1}(r_1) Q_{l_2}(r_2) R^{-l_1-l_2-1}. \quad (18)$$

Here Q_1 is the dipole moment, Q_2 is the quadrupole moment, etc. These coefficients were subtracted from the W -coefficients determined from the Legendre expansion. The R -dependence of the resulting W -coefficients was fitted using the RKHS method with $n_{rkhs} = 2$. The value of m_{rkhs} used depends on the dominant remaining interaction term. For a given interaction term with a dominant asymptotic decay of R^{-n} , m_{rkhs} was chosen to be $n - 1$ to have a correct long range extrapolation.³⁸

The $W_{l_1, m, k_1}^{(1)}(R)$ and $W_{l_2, m, k_2}^{(2)}(R)$ -coefficients with $k_1, k_2 = 1, 2$ were calculated by varying r_1 by -0.3 and $0.3 a_0$ with respect to the equilibrium bond length, while keeping r_2 fixed. For the $\text{Na}_2\text{-K}_2$ arrangements, the diatoms are different, meaning that the reverse was also performed. Their R -dependence was fitted using the same RKHS method but with the m_{rkhs} -values corresponding to the dominant interaction term.

Note that we altered the multipole moments and dispersion coefficients from the literature to fit better to the data. Some coefficients were not available in the literature, and these were inferred directly from the data. It was not possible to do this accurately for the coefficients marked with a star (*) in the table, but in these cases, the actual value of the coefficient was only of minor influence to the final result. In such cases, the coefficients mainly functioned to stabilize the extrapolation of the RKHS fit in the long range.

In the long range, the dominant interaction term in the NaK-NaK arrangement is the dipole-dipole term, with an

TABLE II. The R^{-6} contributions of the W_{000} , W_{200} , and W_{020} coefficients and the multipole moments Q_l at the equilibrium bond length that were inserted into the long-range fitting model. All coefficients are given in atomic units. The values marked with a * had only very minor influence on the fitting result, and their value could therefore not unambiguously be determined.

Parameter	Reference 35	Used value
Q_1 NaK	1.09–1.156	1.068
Q_2 NaK	10.56–10.60	10.6
Q_3 NaK	–26.54	–26.4
Q_4 NaK	...	160
$W_{000}(R^{-6})$ NaK–NaK	7777	8500
$W_{200}(R^{-6})$ NaK–NaK	551.9	700*
Q_2 Na_2	10.52	10.7
Q_4 Na_2	...	73*
Q_2 K_2	15.68	15.9
Q_4 K_2	...	375*
$W_{000}(R^{-6})$ $\text{Na}_2\text{-K}_2$...	7900
$W_{200}(R^{-6})$ $\text{Na}_2\text{-K}_2$...	600*
$W_{020}(R^{-6})$ $\text{Na}_2\text{-K}_2$...	1700*

asymptotic decay of R^{-3} . The dominant interaction between ground state NaK -diatoms in the long range is rotational dispersion, of which the magnitude is determined by the dipole moment.³⁹ The dipole moment is one of the only terms we can accurately determine from our *ab initio* points because it is strongly dominant over the dispersion terms with the same angular dependence at $R = 30a_0$. The dipole moment we extract from our long range fit is given by $1.068 e a_0$, which agrees well with the experimental value⁴⁰ of $1.07 \pm 0.04 e a_0$. There is a difference with the theoretical values found by Byrd ranging from 1.09 to 1.156 $e a_0$,³⁵ where they calculated the dipole moment using a finite field method on the diatom.

E. Merging the short range and the long range

In order to merge the long range with the short range, we first need to match the two *ab initio* methods and then merge the GP model with the analytical long-range part. This was performed in two steps. First the size-consistency error of the CI-calculations was corrected for by matching the long-range and the short-range parts optimally in the parallel and antiparallel NaK-NaK arrangements, by introducing an absolute offset to the CI calculations of -44 cm^{-1} .

In the second step, the long-range energies were calculated from the long-range analytical fit and used as training points for the GP model. The two parts were then merged using the switching function in Eq. (7). The smoothest transition can be made if the change in energy in the merging range of R is bigger than the difference between the two functions to be merged. The long-range fit should only be mixed in past the minima as a function of R for all orientations, i.e., when R is larger than the R -values corresponding to the minima with a collinear orientation. This resulted in the values of $c = 17.5 a_0$ and $w = 2.5 a_0$. This leads to the following modified training points:

$$V_{\text{train}}(\vec{x}_{ab}) = F_{\text{merge}}[R; 17.5, 2.5, V_{\text{ab}}(\vec{x}_{ab}), V_{\text{lr}}(\vec{x}_{ab})]. \quad (19)$$

Then the long-range fits were merged with the short-range fits (trained with the modified *ab initio* points), $V_{\text{Na}_2\text{K}_2}^{\text{GP}}(\vec{x})$, $V_{14}^{\text{GP}}(\vec{x})$, $V_{23}^{\text{GP}}(\vec{x})$ from Eqs. (11)–(13), in the same way,

$$V_{\text{merge}}(\vec{x}) = F_{\text{merge}}[R; 17.5, 2.5, V^{\text{GP}}(\vec{x}), V_{\text{lr}}(\vec{x})]. \quad (20)$$

The resulting expressions enter Eqs. (14) and (15) instead of $V_{\text{Na}_2\text{K}_2}^{\text{GP}}(\vec{x})$, $V_{14}^{\text{GP}}(\vec{x})$, and $V_{23}^{\text{GP}}(\vec{x})$.

Figure 2 shows the result of the procedure described above for a sample orientation. It is clear that the unadjusted GP fit of the CI points does not provide a physically well-behaved long range. The GP fit of the modified points already behaves more physically inside the training point region ($R < 25 a_0$), but this fit starts to increase linearly outside of the training point region. This is due to the dot product kernel in the composite kernel. The final fit obtained by the mixing of the long-range fit, and the GP fit is smooth and physical.

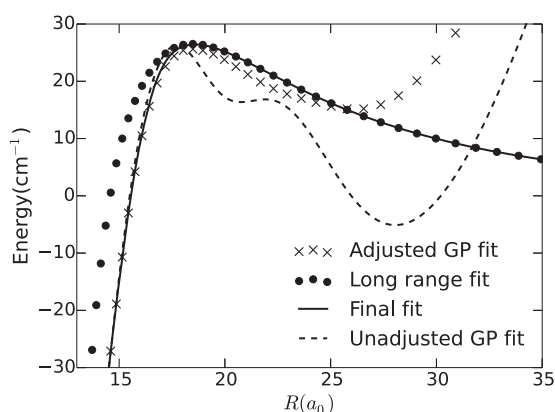


FIG. 2. The transition between the long-range fit and the GP fit in the asymptotic, parallel NaK–NaK arrangement, so with $\theta_1 = \theta_2 = \frac{\pi}{2}$, $\phi = 0$ and at the equilibrium bond lengths.

IV. RESULTS AND DISCUSSION

A. Properties of the PES

In this section, we use the GP model to examine the features of the PES. The geometries were optimized using the GP models in combination with the steepest descent gradient approach. After that, the accuracy of the energy at the minima was improved using geometry optimization at the MRCI level in MOLPRO. The position of the transition states for the GP models was found using a climbing image Nudged Elastic Band (NEB) method.⁴¹ Starting from the transition states found, a transition state search was performed in MOLPRO.

The resulting reaction diagram is shown in Fig. 3. We call two minima symmetrically equivalent if they can be transformed into each other by a permutation of identical atoms but not by an overall, rigid rotation of the complex (in which case the minima would be identical). This means that there is a

transition state between the symmetrically equivalent minima. There is a quartet of symmetrically equivalent minima with the structure indicated by II and a pair of symmetrically equivalent minima with the structure of III. Altogether the NaK–NaK PES therefore has seven minima. Note that the structure of all minima and transition states is planar. Minimum I, of D_{2h} symmetry, has the lowest energy, -4534 cm^{-1} . Minimum II has a C_s symmetry and an energy of -3425 cm^{-1} . The energy of the transition state between minimum I and minimum II, called TS 1 in the diagram, is below the dissociation energy of the complex, meaning that the barrier for the reaction of $\text{NaK} + \text{NaK} \rightarrow \text{Na}_2 + \text{K}_2$ is submerged by 1773 cm^{-1} .

The second lowest local minimum III is reported in the literature³⁵ to also have a D_{2h} symmetry, but here the symmetry is found to be lower: a C_{2h} symmetry. The structure having the D_{2h} symmetry corresponds to the transition state between minimum III and its symmetric equivalent.

There is only one transition state between the symmetrically equivalent minima II shown in Fig. 3, which is the transition between the minima related by \hat{P}_{12} . This transition state is of C_{2v} symmetry. For the other transitions, it is more favorable to first go via TS 1 to minimum I and then back. Similarly, only one of the two transition states between minimum II and minimum III is shown. The other transition state lies higher than TS 3, meaning that it is favorable to first do a rearrangement of minimum II or III to a symmetric equivalent, via TS 2 or TS 4, before going to the other minimum.

The physical reason that minimum III has a C_{2h} instead of D_{2h} symmetry can be explained using simple molecular orbital theory, by looking at the highest occupied molecular orbital (HOMO), which is represented schematically for the three minima in Fig. 4. Here we see that the HOMOs of all minima have a node going through the centre of the complex. We see that the orbitals on the K-atoms mainly have a p -character, meaning that the node of the HOMO goes through the nuclei of the K-atoms. For Na-atoms, the s - p gap is larger than for K-atoms, which means in this case that the p -orbitals have a too high energy to strongly contribute. For minimum III,

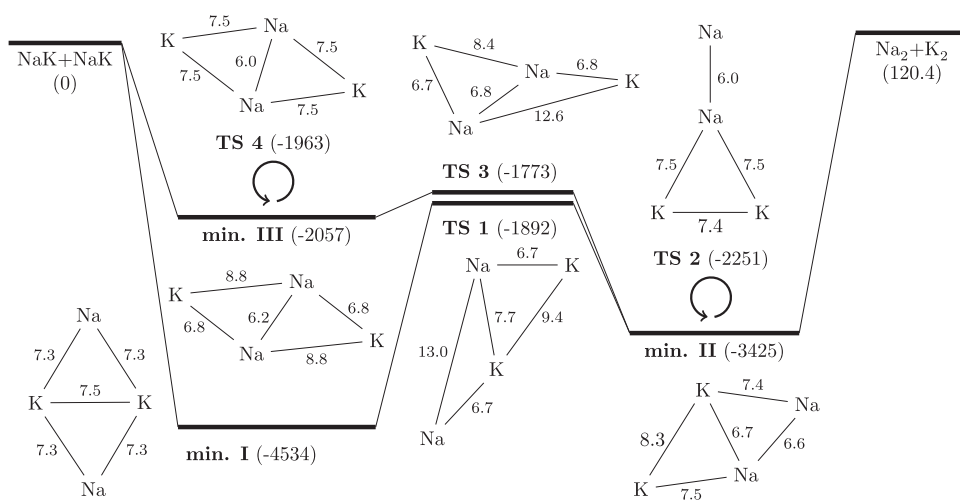


FIG. 3. The geometries and energies (between parentheses, in cm^{-1}) of the minima and the transition states between them. The numbers next to the lines between the atoms are the corresponding bond lengths in a_0 . The structures in all three minima and transition states are planar.

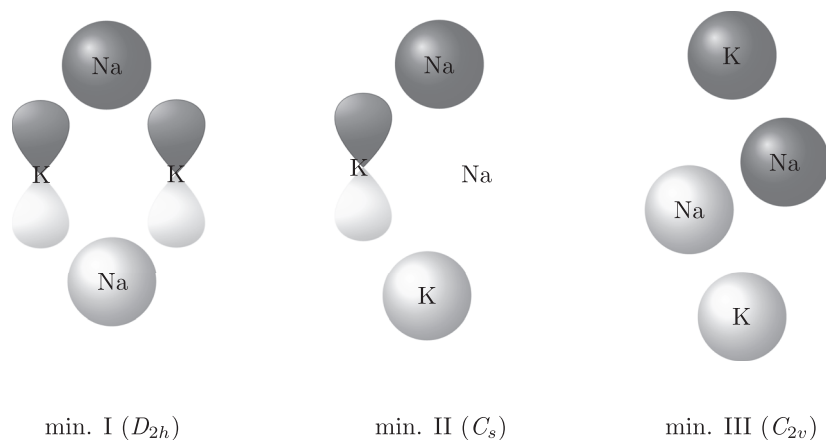


FIG. 4. A schematic representation of the highest occupied molecular orbital of the NaK–NaK complex at the geometries of the minima in Fig. 3.

the orbitals on the Na-atoms, which are central in the complex, have an s-like character but an opposite phase. Each of the two Na-atoms forms a bond with one of the outer K-atoms, but there is an antibonding interaction between the two Na-atoms in the middle. It is thus favorable to break the D_{2h} -symmetry to make the central bond longer and the NaK-bonds shorter. The less favorable interactions in the case of minimum III also lead to a higher energy of the HOMO than in the case of minimum I and therefore also a higher energy of the minimum. In the case of minimum II, there is again a p-orbital on the central K-atom, but here the central Na-atom barely contributes to the HOMO. The energy of the HOMO of minimum II lies in between the HOMO energies of minimum I and minimum III, just as the overall energy of the minimum.

We note that there is also a seam of conical intersection present in the PES. The minima on the conical intersection were sought with MOLPRO on the MC-SCF level without the CPP's (the MOLPRO 2015 optimization algorithm for conical intersections could not handle CPPs). Two of the geometries found are still close to the seam of conical intersections (an energy difference between the ground and excited state of less than 1000 cm^{-1}) with CPPs included. The structure of those minima and their energies are displayed in Fig. 5. The energies at both geometries lie more than 1000 cm^{-1} above the dissociation energy, and we expect this is the case for the entire seam of conical intersections. This also means that the seam

of conical intersections is outside the training point region of the GP. For other alkali tetramers, however, the seam of conical intersections may appear for lower energies and could strongly affect the collision dynamics.

B. Accuracy of the GP models

In total, nine GP models were constructed using the methods outlined in Sec. III, three models with 499–517 (GP1–3), three models with 938–949 points (GP4–6), and three models with 2271–2344 points (GP7–9). The root mean squared errors of these surfaces in the different regions of the potential are listed in Table III.

The error for each GP model was determined by predicting the energy values on the grids used for all other GPs. Table III clearly shows that the error decreases as R increases, i.e., as the interaction becomes weaker. The largest errors in the range from $R = 4$ – 10 rise up to more than 1000 cm^{-1} for points in the repulsive wall. In the classically accessible region, GPs 7–9 provide stable results, with no point where the error

TABLE III. The root mean squared error (in cm^{-1}) and maximum error (in parentheses) of the GP fits [without the second merging step with the long-range fit in Eq. (20)] in different R-ranges of the PES. The number of *ab initio* points N_{ab} used for each fit and the total number of training points N_{tot} (with added symmetrically equivalent points) are also displayed. Column 4–10* displays the errors computed on the subset of points excluding the repulsive wall ($E > 0\text{ cm}^{-1}$).

	N_{ab}	N_{tot}	R-range (in a ₀)			
			4–10	4–10*	10–16	16–25
GP1	499	659	361 (3781)	242 (1564)	84 (951)	12.4 (175)
GP2	512	679	314 (2650)	164 (913)	73 (752)	11.5 (70)
GP3	517	693	396 (3854)	248 (2124)	81 (1280)	11.8 (121)
GP4	938	1257	256 (2378)	156 (1402)	54 (1128)	8.5 (72)
GP5	949	1252	260 (2437)	116 (726)	53 (752)	7.5 (57)
GP6	943	1237	329 (3613)	207 (2292)	73 (1210)	7.7 (52)
GP7	2344	3071	194 (1923)	95 (509)	30 (370)	4.9 (38)
GP8	2274	2970	185 (1803)	96 (659)	36 (585)	4.7 (36)
GP9	2271	2986	170 (1470)	81 (490)	37 (375)	4.7 (33)

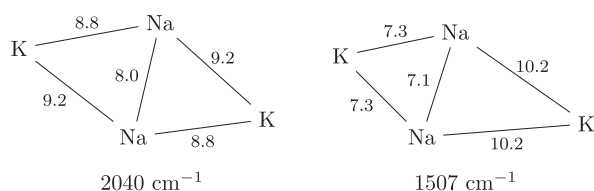


FIG. 5. The geometries corresponding to positions on the seam conical intersections of the NaK–NaK complex; the energies corresponding to these geometries are 2040 and 1507 cm^{-1} . These structures are both planar.

TABLE IV. The minimum energy E_{\min} (in cm^{-1}) and the geometry at the minimum of the GP fits 7-9. We have also denoted the energy of the lowest energy training point and its position. Note that at the minimum, $r_{13} = r_{14} = r_{23} = r_{24}$. This is not the case for the training points, meaning that the range of r_{13} , r_{14} , r_{23} and r_{24} is shown.

	E_{\min}	r_{12}	r_{13}	r_{34}
GP7	-4552	7.39	7.34	12.68
GP7 training set	-3867	7.26	7.31-7.97	12.69
GP8	-4668	7.49	7.33	12.60
GP8 training set	-3700	7.43	6.55-8.30	12.55
GP9	-4594	7.50	7.33	12.61
GP9 training set	-3313	7.42	7.04-8.46	12.00
<i>Ab initio</i>	-4534	7.52	7.35	12.62
Reference 35	-4399	7.76	7.64	12.75

exceeds 700 cm^{-1} . The fits with fewer points have regions where the maximum error is large in the classically accessible region, especially in regions close to minimum III.

The error for the GP models 7-9 is about 40% smaller than the error for the GP models 4-6. The error for different GP models with roughly the same number of points is similar, but the stability improves if more points are used. We estimate that the error of the models GP7-9 is of the same order as the error on the *ab initio* points themselves, which can be estimated by comparing the *ab initio* energies computed here to the literature results (also shown in [Table IV](#)) obtained with a different *ab initio* method. To estimate the effect of the error of the GP regression on calculated observables, a good strategy would be to perform the calculations for three different models of the same PES and evaluate the difference. This illustrates that the effect of the GP error on the collision observables can be easily quantified.

The power of the GP method for interpolation can be further illustrated by analyzing the prediction of the GP models at the global minimum of the PES, shown in [Table IV](#). The prediction accuracy of both the energy and the position of the minimum is remarkable, especially given that the lowest energy points used for training the GP models are $750\text{--}1300 \text{ cm}^{-1}$ higher in energy. This clearly shows that important features of the PES can be captured by GP regression even if the training points are some distance away from these features.

The performance of the GP models for predicting the other minima and the transition states is illustrated in [Table V](#). All GPs predict a structure with the correct symmetry and approximately correct geometry for minimum I and minimum II. The energy is most accurately predicted by GPs 7-9, which is not surprising since those are trained using the most points. The performance of GPs 1-3 is better than the performance of GPs 4-6 for minimum I, while for minimum II, it is the other way around. Minimum III is only well described by the GP fits 5 and 7-9. Since this minimum is shallow and corresponds to a small region in parameter space, the sets of only 500 or 1000 training points do not have enough training points around the minimum to describe it well. The maximum errors from [Table III](#) in the classically accessible part of the PES are also mostly in this region.

The transition states are generally predicted with a similar accuracy as the minima, with an exception of TS 4. This is again a very delicate feature of the PES, and even with 2500 training points, the density of training points is not high enough to accurately cover this transition state. Only for GP5 and GP9, the energy of TS 4 is lower than the energy of TS3. Except in the case of this transition state TS 4, the GP prediction accuracy for GP7-9 is generally 5% or better.

The difference in energy between the calculated *ab initio* points with the literature is mainly due to the difference in the *ab initio* method. In the work by Byrd,³⁵ CCSD(T) is used for this minimum configuration and also the first shell below the valence shell is treated explicitly. Here, a CI method is used and only the valence electrons are treated using ECPs and CPPs. We used a bigger basis for the valence electrons, however. We do not know which method is superior.

Finally, to demonstrate that the GPs indeed provide a smooth fit of the *ab initio* points in the short range, we plotted transition paths from minimum I to minimum II in [Fig. 6](#). All transition paths we plotted have the same start-point and end-point, the minima I and II found from our *ab initio* calculations. The symbols correspond to the nodes of the NEB algorithm. The coordinates at the nodes were interpolated using a quadratic spline, and on the interpolated coordinates, the GPs were evaluated, yielding the lines in the graph. We see that qualitatively, all transition paths are the same and they are all smooth. There are clear differences between the GPs,

TABLE V. The local minimum energies and the transition state energies (in cm^{-1}) of the GP fits 4-6 compared to the *ab initio* and literature energy values. For minimum III, marked with a star (*), Byrd³⁵ reported a structure corresponding to what we find to be transition state 4. GPs 1, 2, 3, and 4 do not contain minimum III, and GP 6 contains a minimum close to minimum III but with C_s instead of C_{2h} symmetry.

	<i>Ab initio</i>	GPs 1-3	GPs 4-6	GPs 7-9	Reference 35
Minimum I	-4534	-4615 ± 59	-4743 ± 161	-4605 ± 60	-4399
Minimum II	-3474	-3256 ± 156	-3645 ± 51	-3559 ± 83	-3113
Minimum III	-2057	...	-2088 (GP5)	-2125 ± 114	-1755*
TS 1	-1892	1820 ± 64	1837 ± 68	-1875 ± 38	...
TS 2	-2251	-2550 ± 312	2143 ± 141	-2172 ± 68	...
TS 3	-1773	...	-1883 (GP5)	1756 ± 10	...
TS 4	-1963	...	-1933 (GP5)	1808 ± 290	...

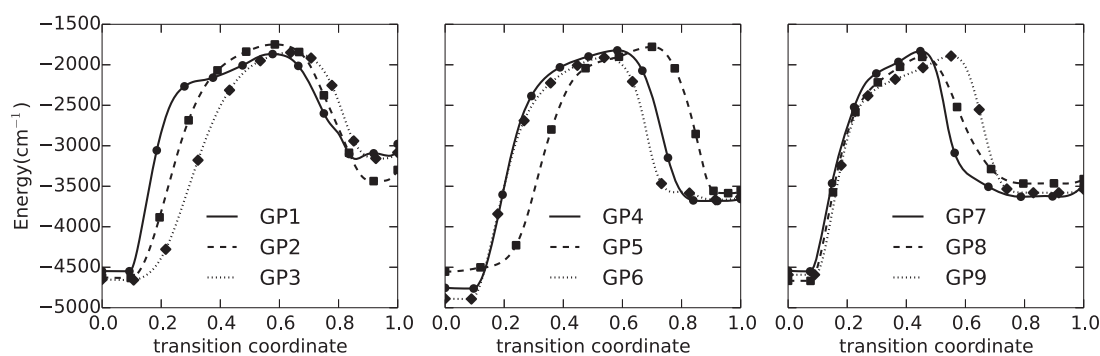


FIG. 6. Transition paths through the transition states between minimum I and minimum II for all the GPs. The begin-point and end-point of all transition paths were the minima I and II of the *ab initio* points. Transition paths were found for each GP using the NEB method⁴¹ with 10 nodes (excluding the begin-point and end-point), marked with symbols. The coordinates of the nodes were interpolated using a quadratic spline, and the interpolated coordinates (1000 points) were evaluated by the GPs to yield the lines in the graphs. The total distance along the transition path was scaled to 1 for each GP individually.

but they become smaller for GPs that are trained using more points, as expected.

V. CONCLUSION

This work presents an accurate potential energy surface in full dimensionality for the NaK–NaK collision complex and the analysis of the features of this reactive system. To the best of our knowledge, this is the first detailed study of the global potential energy surface for interacting polar alkali metal dimers. We have demonstrated a method of merging a machine-learning (GP) model with the analytical forms representing the asymptotic long-range interactions to construct a global surface suitable for scattering calculations. In the [supplementary material](#), we provide the *ab initio* points used to train the GP models and the long-range interaction coefficients. We also provide the codes to construct the GP models from the training sets, the codes to construct the long range fits from the long-range coefficients, and the code to evaluate the final fit consisting of both the short-range and the long-range part (see the [supplementary material](#)). Since Gaussian process regression is a non-parametric statistical learning technique, the method (the codes) presented here should be applicable to any system of four alkali metal atoms.

We have also examined the ability of Gaussian processes to describe the detailed features of the PES. We showed that accurate representations of the global PES can be obtained with 2500 *ab initio* points, an order of magnitude smaller than would be needed for conventional fitting methods. This makes it possible to construct PESs efficiently for the entire range of alkali metal tetramers and for larger systems. We have illustrated that the GP models trained with as few as 500 *ab initio* points capture the prominent features of the global PES. This suggests that GP regression may be an efficient tool for exploring the features of unknown complex PESs. This is particularly important because GP models can be combined with Bayesian optimization¹⁶ that can be used to locate efficiently the minima and saddle points without the need to evaluate the gradient of the potential.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for the training sets, the long range coefficients, and the codes to construct and evaluate our PESs. The training sets already contain the added symmetrically equivalent points (such as mentioned in Sec. III C), and the training point energies have already been modified using the long range energies (such as mentioned in Sec. III E). The unmodified training sets are available upon request. We have also added a document with instructions on how to construct the PESs. This document also contains the hyperparameters of our obtained PESs, which can be used to check whether the PESs have been reproduced correctly.

ACKNOWLEDGMENTS

A.C. would like to thank the Radboud Honours Academy for financial support. The work of R.A.V.-H and R.V.K. was supported by NSERC of Canada.

REFERENCES

- ¹G. Quémener and P. S. Julienne, *Chem. Rev.* **112**, 4949 (2012).
- ²J. L. Bohn, A. M. Rey, and J. Ye, *Science* **357**, 1002 (2017).
- ³J. M. Sage, S. Sainis, T. Bergeman, and D. DeMille, *Phys. Rev. Lett.* **94**, 203001 (2005).
- ⁴J. Deiglmayr, A. Grochola, M. Repp, K. Mörtlbauer, C. Glück, J. Lange, O. Dulieu, R. Wester, and M. Weidemüller, *Phys. Rev. Lett.* **101**, 133004 (2008).
- ⁵K.-K. Ni, S. Ospelkaus, M. H. G. de Miranda, A. Pe'er, B. Neyenhuis, J. J. Zirbel, S. Kotochigova, P. S. Julienne, D. S. Jin, and J. Ye, *Science* **322**, 231 (2008).
- ⁶T. Takekoshi, L. Reichsöllner, A. Schindewolf, J. M. Hutson, C. R. L. Sueur, O. Dulieu, F. Ferlaino, R. Grimm, and H.-C. Nägerl, *Phys. Rev. Lett.* **113**, 205301 (2014).
- ⁷P. K. Molony, P. D. Gregory, Z. Ji, B. Lu, M. P. Köppinger, C. R. L. Sueur, C. L. Blackley, J. M. Hutson, and S. L. Cornish, *Phys. Rev. Lett.* **113**, 255301 (2014).
- ⁸J. W. Park, S. A. Will, and M. W. Zwierlein, *Phys. Rev. Lett.* **114**, 205302 (2015).
- ⁹M. Guo, B. Zhu, B. Lu, X. Ye, F. Wang, R. Vexiau, N. Bouloufa-Maafa, G. Quémener, O. Dulieu, and D. Wang, *Phys. Rev. Lett.* **116**, 205303 (2016).
- ¹⁰P. S. Żuchowski and J. M. Hutson, *Phys. Rev. A* **81**, 060703 (2010).

- ¹¹M. Mayle, G. Quémener, B. Ruzic, and J. L. Bohn, *Phys. Rev. A* **87**, 012709 (2013).
- ¹²X. Ye, M. Guo, M. L. González-Martínez, G. Quémener, and D. Wang, *Sci. Adv.* **4**, eaaq0083 (2018).
- ¹³C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Massachusetts Institute of Technology, 2006).
- ¹⁴J. Cui and R. V. Krems, *J. Phys. B: At., Mol. Opt. Phys.* **49**, 224001 (2016).
- ¹⁵B. Kolb, P. Marshall, B. Zhao, B. Jiang, and H. Guo, *J. Phys. Chem. A* **121**, 2552–2557 (2017).
- ¹⁶R. A. Vargas-Hernández, Y. Guan, D. H. Zhang, and R. V. Krems, “Bayesian optimization for the inverse scattering problem in quantum reaction dynamics,” *New J. Phys.* (to be published); preprint [arXiv:1711.06376](https://arxiv.org/abs/1711.06376) (2018).
- ¹⁷A. Kamath, R. A. Vargas-Hernández, R. V. Krems, T. Carrington, and S. Manzhos, *J. Chem. Phys.* **148**, 241702 (2018).
- ¹⁸C. Qu, Q. Yu, B. L. Van Hoozen, J. M. Bowman, and R. A. Vargas-Hernández, *J. Chem. Theory Comput.* **14**, 3381 (2018).
- ¹⁹H.-J. Werner, P. J. Knowles et al., *MOLPRO*, version 2012.1, a package of *ab initio* programs, 2012, see <http://www.molpro.net>.
- ²⁰A. Nicklass, M. Dolg, H. Stoll, and H. Preuss, *J. Chem. Phys.* **102**, 8942–8952 (1995).
- ²¹A. Nicklass and K. A. Peterson, “Core-valence correlation effects for molecules containing first-row atoms. Accurate results using effective core polarization potentials,” *Theor. Chem. Acc.* **100**, 103 (1998).
- ²²P. Fuentealba, H. Preuss, H. Stoll, and L. V. Szentpály, *Chem. Phys. Lett.* **89**, 418 (1982).
- ²³A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.6.1), Online. Available: <https://physics.nist.gov/asd>; 22 November 2018, National Institute of Standards and Technology, Gaithersburg, MD, 2018.
- ²⁴A. Gerdes, M. Hobein, H. Knöckel, and E. Tiemann, *Eur. Phys. J. D* **49**, 67–73 (2008).
- ²⁵R. Ferber, E. A. Pazyuk, A. V. Stolyarov, A. Zaitsevskii, P. Kowalczyk, H. Chen, H. Wang, and W. C. Stwalley, *J. Chem. Phys.* **112**, 5740 (2000).
- ²⁶J. A. Pople, R. Seeger, and R. Krishnan, *Int. J. Quantum Chem.* **12**, 149 (1977).
- ²⁷S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.* **8**, 61 (1974).
- ²⁸S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
- ²⁹D. Duvenaud, H. Nickisch, and C. Rasmussen, *Adv. Neural Inf. Process. Syst.* **24**, 226 (2011).
- ³⁰D. Duvenaud, J. Lloyd, R. Grosse, J. Tenenbaum, and G. Zoubin, in *Proceedings of the 30th International Conference on Machine Learning Research (JMLR.org, 2013)*, Vol. 28, pp. 1166–1174.
- ³¹R. A. Vargas-Hernández, J. Sous, M. Berciu, and R. Krems, *Phys. Rev. Lett.* **121**, 255702 (2018).
- ³²F. 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, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- ³³M. Stein, *Technometrics* **29**, 143–151 (1987).
- ³⁴F. O. Ellison, *J. Am. Chem. Soc.* **85**, 3540 (1963).
- ³⁵J. Byrd, “Ultracold chemistry of alkali clusters,” Ph.D. thesis, University of Connecticut, 2013.
- ³⁶T.-S. Ho and H. Rabitz, *J. Chem. Phys.* **104**, 2584 (1996).
- ³⁷M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C., 1964).
- ³⁸T.-S. Ho and H. Rabitz, *J. Chem. Phys.* **113**, 3960 (2000).
- ³⁹P. S. Żuchowski, M. Kosicki, M. Kodrycka, and P. Soldán, *Phys. Rev. A* **87**, 022706 (2013).
- ⁴⁰R. F. Wormsbecher, M. M. Hessel, and F. J. Lovas, *J. Chem. Phys.* **74**, 6983 (1981).
- ⁴¹G. Henkelman, B. Uberuaga, and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000).