# Automating the Development of High-Dimensional Reactive Potential Energy Surfaces with the ROBOSURFER Program System

Tibor Győri* and Gábor Czakó*

MTA-SZTE Lendület Computational Reaction Dynamics Research Group, Interdisciplinary Excellence Centre and Department of Physical Chemistry and Materials Science, Institute of Chemistry, University of Szeged, Rerrich Béla tér 1, Szeged H-6720, Hungary

S *Supporting Information*

ABSTRACT: The construction of high-dimensional global potential energy surfaces (PESs) from *ab initio* data has been a major challenge for decades. Advances in computer hardware, electronic structure theory, and PES fitting methods have greatly alleviated many challenges in PES construction, but building fitting sets has remained a bottleneck so far. We present the ROBOSURFER program system that completely automates the generation of new geometries, performs *ab initio* computations, and iteratively improves the PES under development. Unlike previous efforts to automate PES development, ROBOSURFER does not require any uncertainty estimate from the PES fitting method and thus it is compatible with the permutationally invariant polynomial (PIP) method. As a demonstration we have developed five related but different global reactive PIP PESs for the $CH_3Br + F^-$ system and used them to perform quasiclassical trajectory (QCT) reaction dynamics simulations over a wide range of collision energies. The automatically developed PESs show good to excellent accuracy at known stationary points without any manual sampling, and QCT results indicate the lack of unphysical minima on the fitted surfaces. We also present evidence suggesting that the breakdown of single reference electronic structure theory may contribute significantly to the fitting errors of global reactive PESs.

Introducing the ROBOSURFER program system
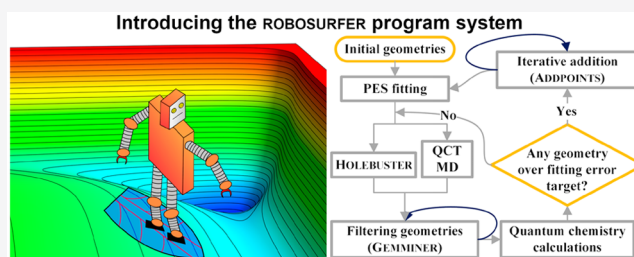
I. INTRODUCTION

The reactive collision of atoms and molecules has been a cornerstone of chemical kinetics ever since the proposal of collision theory[1] in the early 20th century. Likewise, an understanding of elementary bimolecular reactions, such as bimolecular nucleophilic substitution (S_N2), has been indispensable for the elucidation of more complicated reaction mechanisms.

In the last couple of decades, the simulation of reactive molecular collisions has become increasingly feasible and fruitful, resulting in new insights[2] into previously known reaction mechanisms[3] and the discovery of entirely new ones.[2,4] This has been enabled not only by the explosive growth of available computing power but also by breakthroughs in fitting analytical potential energy surfaces (PESs)[5−10] to high quality *ab initio* data.

The motion of nuclei within the framework of the Born–Oppenheimer approximation[11] can be simulated using either quantum or (quasi)classical trajectory (QCT)[12] methods. In both cases, a large number of potential energy (and possibly energy gradient) evaluations are required for the calculation of statistically robust integral and differential cross sections. For this reason, on-the-fly computation of energies and gradients via quantum chemical methods (the so-called direct dynamics approach) is only feasible using DFT and MP2 methods, with modest basis sets, thus potentially impacting the accuracy of such studies.[13]

An alternative of direct dynamics is the construction of an analytical potential energy surface, where a limited number of high quality energies (and possibly energy derivatives) are calculated at a diverse set of molecular geometries, and then a suitable procedure is used to fit a function that is orders of magnitude faster to evaluate than the original quantum chemical (QC) method. Browsing through the literature reveals a large number of proposed fitting procedures, including the following: various modified Shepard interpolation (MSI) schemes,[6,14] interpolating moving least-squares (IMLS),[15,16] fitting permutationally invariant polynomials (PIP),[5,17−19] and methods based on Gaussian Process Regression (GP),[20,21] as well as various methods using artificial neural networks (NN)[22−24] and combinations of PIP and the latter two methods (PIP-GP[25] and PIP-NN[26,27]).

While the benefits of using high quality *ab initio* energies are clear,[13,28] constructing a high-dimensional PES with globally low fitting error is nontrivial, since beyond the choice and tuning of the fitting procedure, one also has to select the geometries used for the fit. A uniformly dense, grid based sampling of the configuration space is simple but quickly becomes unfeasible for larger systems due to the exponential growth of the number of samples required.

Commonly employed sparse sampling methods include generating systematic or random displacements along the normal modes of the reactants, known stationary points and known products,[5,6,19] sampling along known minimum energy paths (MEPs),[6,14,19] Sobol sequences,[16] running direct classical dynamics,[5,19] and running classical[6,14] or quantum dynamics[29] using a preliminary PES.

While these methods (or a combination of them) can provide reasonably accurate PESs that are suitable for QCT and quantum dynamics simulations,[5,18,30−35] they suffer from a number of downsides. Systematic approaches that rely on chemical intuition may fail to adequately sample regions important for yet-to-be-discovered mechanisms or unexpected product channels. It is well-known that reaction dynamics often do not follow MEPs[36] and may explore regions far from them, especially at higher collision energies.

Although running trajectories and saving geometries at every Nth time step of the dynamics simulations certainly is a suitable approach for generating samples in chemically interesting regions, on its own it is not a perfect solution for constructing a fitting set.

First, if one uses an affordable direct dynamics method or a preliminary PES, then one might get a different distribution of geometries compared to the hypothetical case of running direct dynamics using the chosen high quality *ab initio* method. Using direct dynamics may have another drawback when used for developing reactive PESs: it may not be feasible to run enough trajectories to ensure that even low probability product channels are adequately sampled.

Second, the set of geometries generated this way is excessively large and often largely consists of reactants approaching and products separating. Randomized and Sobol-sequence-based sampling is also prone to generating points that end up being redundant or in a very high energy region.

Minimizing the number of geometries necessary for the construction of the PES is clearly desirable. Aside from the cost of having to compute more *ab initio* energies, including unnecessary geometries in the PES can have detrimental effects on its quality or usability. The cost of evaluating interpolative PESs typically increases with the number of geometries used in the PES, and for noninterpolative PESs adding geometries needlessly to the fitting set may result in poor global accuracy in regions of less dense sampling.

A program capable of taking a large number of geometries generated by an arbitrary method and selecting only the ones most likely to improve the PES would certainly aid in the rapid development of new PESs. Furthermore, such a program could also be integrated with the PES fitting and geometry generation routines, yielding a program system that completely automates the construction and validation of high-dimensional PESs.

The idea of automatic iterative PES development is not unprecedented; in fact Ischtwan and Collins[14] have described such a method as early as 1994, using classical molecular dynamics (MD) to generate points and a measure of geometric distance to select the geometries that need to be added to the PES, with the idea being that the fitting error is more likely to be high at points which are far from all points in the fitting set. The GROW package[6] of Collins et al. was perhaps the first example of an integrated PES development tool. While their early implementations appear to have used a simple distance-based selector for adding new geometries, later versions seem to have relied on more sophisticated measures of uncertainty, derived

from the variance of the Taylor expansions used in the interpolative PES.[6]

Unfortunately, the MSI method they have used has a number of drawbacks: the computational cost of getting a single-point energy increases with the size of the fitting set, and it requires the first and second derivatives of the energy at every point in the fitting set. The latter is especially ruinous, considering that for many state-of-the-art quantum chemical methods (CCSD(T)-F12,[37] QVCCD(T),[38] DLPNO-CCSD(T),[39,40] to name a few) even first derivatives are either unavailable analytically or very costly to calculate.

Majumder et al. recently discussed[16] an automated PES development scheme based on the IMLS interpolation method that uses the squared difference between a lower and a higher order polynomial fit to estimate the uncertainty of fit at arbitrary geometries and adds geometries where the uncertainty is maximal. A completely automated, nearly black-box PES development package (AUTOSURF) that uses this technique has very recently been published by some of the authors of ref 16.[41] This implementation is currently limited to treating van der Waals systems of two rigid fragments, and the automated selection of new fitting points is intimately tied to the IMLS interpolation method, preventing a simple replacement of the fitting procedure, while reusing the automated PES development routines.

A number of other publications mention more or less automated PES development tools or schemes for constructing GP and NN PESs. GP PESs can directly provide a statistically well founded uncertainty estimate at arbitrary points,[20] and for NN PESs one can use a committee of NNs trained on the same fitting set but with different initial coefficients[42] to derive a measure of uncertainty from the degree of disagreement between them.

During the preparation of this article, Abbott et al. published the PES-Learn software package[43] that appears to be able to produce accurate NN and GP PESs completely automatically, without iterative fitting set generation. The authors developed PESs for H_2O, H_3O^+, $OCHCO^+$, and H_2CO and evaluated the accuracy of the PESs primarily though fitting errors and by running vibrational calculations. It remains to be seen if this approach is suitable for creating high-dimensional reactive PESs for reaction dynamics purposes.

After this brief survey of literature, it appears that almost all examples of automated PES development tools rely on the uncertainty supplied by the fitting method to choose new geometries. This tight integration, however, precludes the use of these tools for developing PESs with fitting methods that cannot provide uncertainty estimates. Most notably this includes the PIP method that has seen widespread use in spectroscopic,[44] classical dynamics,[18,45] and quantum dynamics[31,35] applications, presumably due to its ease of use, great accuracy, applicability to extended systems as large as the formic acid dimer,[44] and low cost of evaluating the fitted function.

To address the shortcomings of the traditional techniques of the fitting set generation we have developed the ROBOSURFER program system that integrates the PIP PES fitting and QCT programs previously used by our group and several new pieces of software. It incorporates novel techniques for generating and selecting geometries that are likely to improve the PES without requiring any uncertainty feedback from the PES fitting method. This latter feature not only makes it suitable for developing PIP PESs but also enables a modular implementation that could
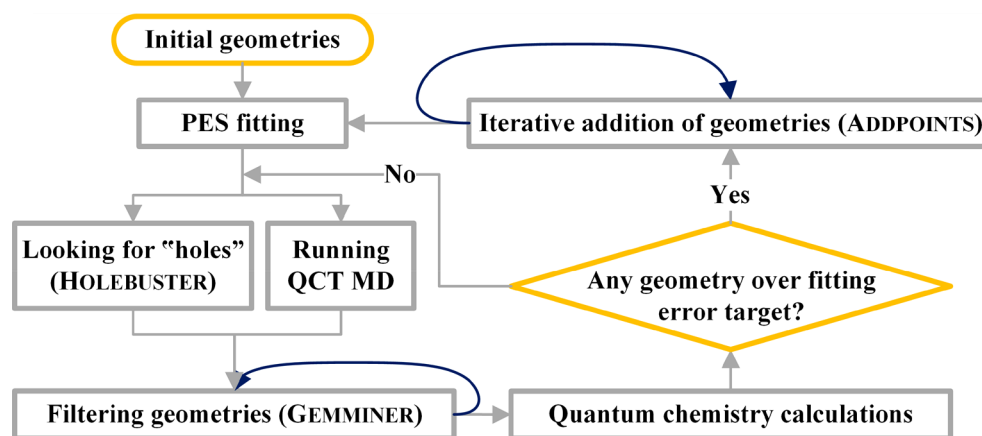
Figure 1. Simplified operational flowchart of the ROBOSURFER program system.

easily be modified to work with any fitting method or extended with other methods of geometry generation.

As a demonstration, we have developed five (related but different) full-dimensional reactive PESs for the $CH_3Br + F^-$ system and used them to perform QCT reaction dynamics simulations over a wide range of collision energies.

The paper is arranged as follows: Section II details the structure and operation of the ROBOSURFER program system and its components, Section III presents a rebuilding scheme, Section IV contains computational details pertaining to the developed PESs and the QCT results, Section V discusses the properties and accuracy of the PESs and presents the QCT results, and finally Section VI gives a condensed summary of the results and conclusions and an outlook on further research.

II. ROBOSURFER PROGRAM SYSTEM

II.A. Overview. The ROBOSURFER program system aims to fully automate the construction and validation of high-dimensional PESs and is designed to be fitting procedure agnostic, highly modular, parallelizable, and easily extensible. Most steps seen in Figure 1 are already able to take advantage of shared-memory parallelism. All subprograms in the ROBOSURFER program system are separate executables and are started by the driver program that maintains the geometry sets. Further implementation details for every subprogram are available in the SI.

The system relies on an iterative improvement approach (Figure 1), where an initial set of geometries are fitted, new geometries are generated and filtered using the fit, QC calculations are run at the filtered geometries, and geometries over the fitting error target (E_{thresh}) are selectively added to the fitting set. In our current implementation this iterative sampling of the configurational space is continued until the user interrupts the program. The initial geometries may be generated *de novo* by any combination of the traditional methods or selected from the fitting set of a similar PES using the later described rebuilding procedure. While *a priori* knowledge of reaction paths and stationary points is useful for the generation of the initial geometries, beyond that, the system does not currently use any information about them.

The core idea behind the ROBOSURFER package is that the overall quality of the fitted PES is best improved by adding points to the fitting set where the fitting error is the highest, given the current fit. It might be worth noting that this concept is very similar to the concept of active learning in machine learning; in fact, one could say that ROBOSURFER is an implementation that builds PESs by a form of active learning.

The exact evaluation of the fitting error is of course not feasible for large numbers of points, since it requires performing the QC calculation. To make active learning feasible with PES fitting methods that cannot supply an uncertainty prediction for arbitrary points, we use geometric and energetic similarity as a rough approximation of the fitting error of new geometries, and the selection is performed by the GEMMINER subprogram.

The driver program maintains four sets of geometries on disk: the ones used for fitting, spares, rejects, and crashes. Spares are geometries for which the QC calculation had been run without error and no rejection criteria was met; these are only excluded from the fitting set due to their (currently) low scaled fitting error (defined later in Section II.G). Rejects are geometries that have been rejected after completing their QC calculations due to being outside the allowed energy window. Crashes are geometries for which the QC calculation has terminated abnormally, typically due to convergence issues.

In the filtering stage, the driver program runs the GEMMINER subprogram four times, first to select the most promising geometries from the newly generated ones and then three more times to discard any new geometries that are too similar to any member of spares, crashes, or rejects, to avoid needless QC calculations.

When the final set of new geometries is ready, the driver program generates input files for the QC package and launches the computations. After all instances of the QC program terminate, the driver program checks and reads the outputs. Problematic geometries are appended to rejects or crashes, and geometries that pass all checks are appended to spares. If any of the latter geometries had a fitting error larger than the fitting error target (E_{thresh}) configured by the user, the ADDPOINTS subprogram is started; otherwise, the current PES is reused and the cycle restarts from the generation of new points (as seen in Figure 1).

The observant reader may note that so far no new geometries have been added to the fitting set. To minimize the number of fitting points, not every new point with an acceptable QC energy is added to the fitting set; regulating this is the responsibility of the ADDPOINTS subprogram. ADDPOINTS calculates the later defined scaled fitting error (SFE) for all geometries in the spares set (including the new geometries just added), moves some geometries with a high SFE to the fitting set, redoes the fit, and recalculates the SFEs of the remaining spares. This iterative

transfer continues until the highest SFE in the set of spares is lower than the E_{thresh} configured by the user.

II.B. PES Fitting. The ROBOSURFER package currently supports two different implementations of the PIP fitting method, one based on the theory of polynomial invariants,[5] and one based on monomial symmetrization,[17] the latter of which is publicly available as the source code.[46] Both implementations lead to the same results, albeit with different trade-offs of code generation complexity and function evaluation cost. The fitting process is mostly a linear least-squares (LLS) solve, although the size of the matrix involved might pose a problem for high-order fits or large systems. Parallelization is achieved by using a parallelized LLS solver. More details on the LLS solvers used can be found in the SI.

II.C. HOLEBUSTER Subprogram. The problem of having deep unphysical minima on PESs has been noted in the literature a number of times.[47−49] In our experience, the presence of these "holes" is a typical symptom of not having enough fitting points in a given PES region.

These artifacts are highly detrimental to any dynamics simulation. In the case of QCT, trajectories entering the hole encounter extreme gradients, causing nuclear motions that are too fast to be accurately handled by the MD integrator at the given time step, resulting in the violation of energy conservation and the formation of thermodynamically impossible products.

Holes are also very problematic for quantum dynamics studies, especially since traditional quantum dynamics methods often evaluate the PES at regularly spaced grid points, without any regard for what regions are accessible by classical dynamics. Therefore, the elimination of spurious minima is required even in high energy regions, if one desires a PES well suited for quantum dynamics.

The HOLEBUSTER subprogram is a modified geometry optimization program that uses a combination of Newton's method and random displacements to find minima. HOLEBUSTER instances are spawned by the driver program, every instance starts from a geometry randomly selected from the current fitting set. Instances are independent and can be run in parallel. All optimization steps are saved, concatenated, and passed along to the GEMMINER subprogram.

II.D. Running QCT Dynamics. QCT dynamics simulations are performed using the latest fit of the PES being developed. The user can set the number, the maximum length, and the collision energy of the trajectories, as well as the set of impact parameters being used. The driver program has provisions for automatically adjusting the length limit of the trajectories, if desired. The current geometry is saved every Nth time step, where N is user-configurable and typically between 2 and 10; these geometries are passed along to the GEMMINER subprogram.

II.E. GEMMINER Subprogram. The overall goal of the GEMMINER subprogram is to take a large set of freshly generated geometries and select a small subset of geometries that have the highest likelihood of having large fitting errors and, thus, the highest likelihood of improving the fitted PES, were they to be added to the fitting set. The GEMMINER subprogram is also essential for the avoidance of geometries where the electronic structure method used in the QC calculations runs into trouble, such as the nonconvergence of the Hartree−Fock iterations.

One of the core ideas behind the GEMMINER subprogram is to use a measure of similarity to the fitting set as a rough approximation of the uncertainty of fit at an arbitrary geometry. This intuitively means that a geometry very close to one of the geometries already used for fitting has a high likelihood of having a low fitting error, while a geometry far away from all fitting points is more likely to have a high fitting error (a rigorous proof of this is beyond the scope of this work). To enhance PES development, GEMMINER may also recommend a geometry based on energy considerations, as described later in this section.

For quantifying geometric similarity, one needs to choose a combination of a coordinate system and a distance function. Ideally, the combination should yield distances that are invariant not only under the translation or rotation of any of the geometries being compared but also the permutation of like atoms.

Translational and rotational symmetry can easily be ensured by using internuclear distances (r_{ij}) as the coordinate system. When comparing a pair of N-atom geometries A and B (with $\mathbf{R}_A$ and $\mathbf{R}_B$ denoting their distance matrix representations), one could use the root-mean-square difference (RMSD) of the corresponding internuclear distances (eq 1) as the distance function,

$$D_{AB}(\mathbf{R}_A, \mathbf{R}_B) = \sqrt{\frac{1}{\frac{N*(N-1)}{2}} \sum_{i=1}^{N} \sum_{j=1}^{i-1} (r_{A,ij} - r_{B,ij})^2} \tag{1}$$

but it has two major shortcomings: the distance (D_{AB}) is overly sensitive to small differences in large internuclear distances, and it is not permutationally invariant.

The first problem results in an unnecessarily dense sampling of the asymptotic regions of the PES: it is easy to understand intuitively that a 0.1 Å difference in the length of a C−F bond is much more significant at a C−F distance of 1.5 Å than at 20 Å. This problem can be solved by replacing the internuclear distances with Morse-like variables (eq 2), yielding the Exponentially Weighted (EW) RMSD metric.

$$D_{AB}(\mathbf{Y}_A, \mathbf{Y}_B) = \sqrt{\frac{1}{\frac{N*(N-1)}{2}} \sum_{i=1}^{N} \sum_{j=1}^{i-1} (y_{A,ij} - y_{B,ij})^2}$$
$$\text{where } y_{ij} = e^{-r_{ij}/a} \tag{2}$$

The second problem can result in spuriously large distances between permutationally equivalent geometries, reducing the efficacy of geometry filtering. We solve this issue by generating the set of all permutations of geometry A (denoted as P), calculating their EW-RMSD distance to geometry B, and returning the minimum distance, yielding the novel Permutationally Invariant (PI) EW-RMSD metric (eq 3).

$$D_{AB}^* = \min_{Q \in P} D_{AB}(\mathbf{Y}_Q, \mathbf{Y}_B) \tag{3}$$

Generating every permutation is nontrivial for systems with large permutational symmetry. For a chemical system with K different atom types where K_i is the number of atoms in the ith atom type, the number of permutations is

$$p = \prod_{i=1}^{K} K_i! \tag{4}$$

To minimize the computational burden and the implementation effort, we created a simple code generator program that uses Heap's algorithm[50] to generate code that only requires one pairwise swap per permutation to generate every permutation in a given group. Using this technique, we have implemented the PI-EW-RMSD metric for systems as large as X_6Y_2Z, yielding 1440 permutations.
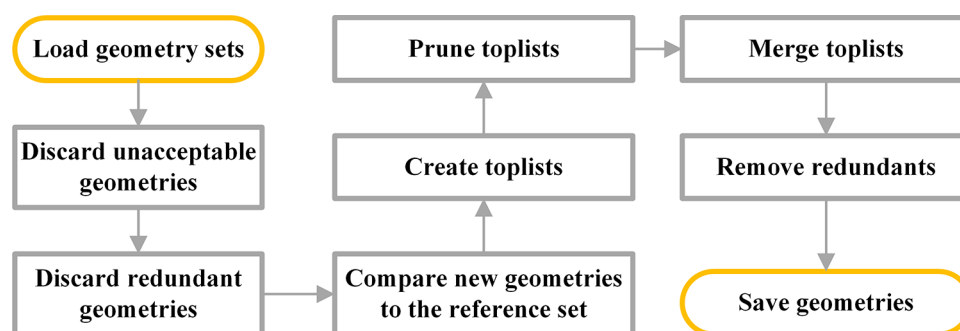
Figure 2. Simplified operational flowchart of the GEMMINER subprogram.

GEMMINER starts by loading two sets of geometries, the reference set and the set of geometries that are to be filtered (for brevity the latter will be referred to as "new geometries" in this subsection) (Figure 2).

New geometries are checked for violations of energy and coordinate constraints, and geometries outside the user configured energy window and Cartesian coordinate limits are discarded. These two constraints are typically set to very permissive values, such as an energy window of ±500 kcal/mol relative to reactants and a coordinate limit of ±200 Å, and are generally only meant to discard the most nonsensical geometries. It should be noted that during all GEMMINER runs the energy associated with new geometries is their energy on the current PES.

Then, new geometries are compared to each other using the EW-RMSD distance metric, and any geometry that has a neighbor closer than a user configurable threshold is discarded. This step is purely done for the sake of saving CPU time, as MD trajectories can supply on the order of 10^5 geometries and would bog down the next stage. More details regarding the performance trade-off of performing this pruning step can be found in the SI.

The pruned new geometries are then compared to every reference geometry using the PI-EW-RMSD distance metric, and for every new geometry the following is saved:

1. Its distance to the closest point in the reference set
2. The absolute value of the energy difference between it and the closest point in the reference set (E_{diff})
3. The energy difference between it and the lowest energy point in the reference set ($E_{\text{diff(GM)}}$)

The first value is saved to facilitate the sampling of regions where the reference set is sparse, the second is to enhance the sampling of regions of high gradient, and the last one is to promote hole fixing. This comparison step is parallelized via OpenMP, and it is often the rate limiting step in GEMMINER.

After this, toplists are constructed using these three values; for the first two the lists are sorted in descending order, and the last one is sorted in ascending order. The toplists are then pruned to enforce a number of constraints that prevent the recommendation of geometries wildly different from the reference set. This was necessary to implement, as sometimes HOLEBUSTER was a little too effective at finding holes, so much so that all geometries recommended by GEMMINER were in extreme regions and suffered from self-consistent field (SCF) nonconvergence. These constraints are user configurable and are turned off by the driver program when the reference set is not the fitting set.

In the case of the first and second toplists, entries that are too close to a reference geometry are also removed. In addition to this, entries in the second toplist with a low E_{diff} are also removed. Finally, geometries in the third toplist are removed if their $E_{\text{diff(GM)}}$ is greater than $-E_{\text{thresh}}$.

After this round of pruning the top M entries in each toplist are concatenated, where M is a user-configurable value, resulting in a new list containing at most $3M$ geometries. Any possible duplicates and near-duplicates are removed from this list by comparing the list members to each other using the PI-EW-RMSD distance metric. Finally, all geometries that remain in the combined list are written to disk.

The driver program runs GEMMINER four times during the filtering stage. The first run uses the fitting set as the reference set and the concatenated geometry outputs of the HOLEBUSTER and QCT programs as new geometries. The subsequent three runs all use the geometries written by the run before them as new geometries, and use the spares, rejects, and crashes as reference sets. The first run has all constraints enabled, while subsequent runs disable the use of constraints and energy-based toplists, their only purpose being the removal of geometries too similar to any geometry found in spares, rejects, or crashes.

The output of the fourth run is the final recommended geometry set which is used in the next step for generating QC inputs. In our experience, this set typically contains 5−60 geometries, if M is set between 20 and 100.

II.F. Running Quantum Chemical Calculations. Input files for the QC calculations are generated automatically, based on a simple template. While the current implementation only supports the use of MOLPRO[51,52] as a QC backend, it would certainly be possible to add support for other QC packages as well. The driver program can launch multiple instances of the QC program in parallel and monitor their return value to determine if a computation has terminated normally.

If one is performing the PES development at a somewhat expensive level of electronic structure theory (such as MP2-F12 with a triple-ζ basis set), it is possible to do a preliminary calculation to get a rough estimate of the energy and abort any calculation that is clearly in a very high energy region. For example, if a HF calculation with a double-ζ basis set shows that the energy is very much higher than the top end of the energy window set for the development of the PES, then the subsequent HF and MP2-F12 calculations with a larger basis set can be safely avoided, and the geometry can be added to rejects. This technique has the added benefit of improving SCF convergence, as smaller basis sets often have better convergence properties, and the converged small-basis wave function can serve as a good guess for the later SCF iterations.

After all instances of the QC program terminate, the driver program checks and reads the outputs. All geometries that pass the checks are appended to spares.
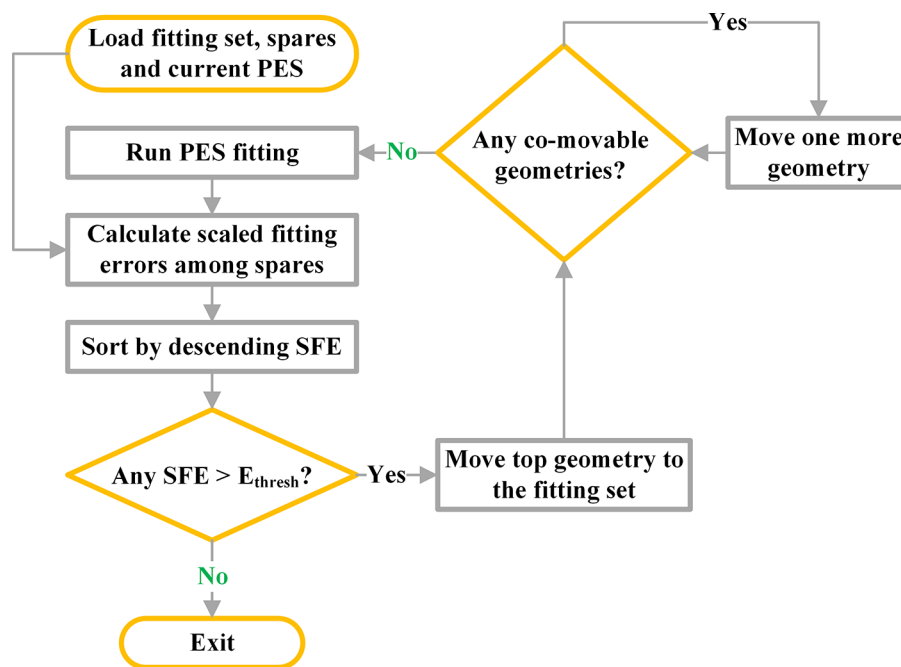
Figure 3. Simplified operational flowchart of the ADDPOINTS subprogram.

II.G. ADDPOINTS Subprogram. The purpose of ADDPOINTS is to selectively move geometries into the fitting set from spares, in an effort to lower the scaled fitting error (SFE) of all geometries in spares below the fitting error target (E_{thresh}), while keeping the fitting set as small as possible. The SFE of a geometry is defined as the product of its absolute fitting error and the scaling factor defined by eq 5:

$$f(E_{pot}) = \begin{cases} \dfrac{E_{thresh}}{E_{thresh} + E_{pot} - E_{pot(max)}}, & \text{if } E_{pot} > E_{pot(max)} + E_{thresh} \\ 1, & \text{if } E_{pot} \leq E_{pot(max)} + E_{thresh} \end{cases}$$
$$(5)$$

where E_{pot} is the energy of the geometry calculated by the QC program and $E_{pot(max)}$ is the maximum energy at which the user wishes to get full accuracy from the PES being developed. The form of f ensures that high energy geometries are less likely to be included in the fit.

As Figure 3 shows, ADDPOINTS begins by loading the fitting set, spares, and current PES coefficients. Then, the scaled fitting error is calculated for all spares and spares are sorted by descending SFE. If the top entry has a SFE lower than E_{thresh}, ADDPOINTS quits; otherwise the top entry is flagged to be moved into the fitting set, and the program enters the geometry addition loop.

While adding geometries to the fitting set one at a time is the most precise scheme, it is not computationally economical, as it requires refitting the PES and recalculating all SFEs after each added geometry. The alternative is moving multiple geometries into the fitting set at the same time; this, however can result in suboptimal results: it is possible that adding just one of them to the fit would change the PES so much that all of the other points would no longer qualify for addition. This is especially common if most new geometries originated from HOLEBUSTER. Thus, naïvely comoving geometries may compromise the goal of minimizing the number of geometries used in the fitting set.

To resolve this conflict between speed and accuracy, we only comove geometries that differ in QC energy by at least 20 kcal/

mol. This heuristic is by no means exact, but it provides some defense against comoving too many geometries from a given PES region.

After the geometry addition loop finishes moving geometries, the PES fitter program is started to obtain the new PES coefficients and the new SFEs of the remaining spares are calculated using the new PES (Figure 3). The cycle of adding a few points and refitting continues until the largest SFE among the spares falls below E_{thresh} or there are no more spares left.

III. REBUILDING PROCEDURE

The ADDPOINTS subprogram of the ROBOSURFER package serves a second purpose: it can also be used standalone to take a large and diverse set of geometries with already computed QC energies and select a subset of them that yields a PES with an SFE no more than E_{thresh} at the excluded geometries. We call this a "rebuild", and it is performed as the following:

1. The initial fitting set and spares are prepared. The fitting set is populated with one or a handful of chosen geometries; all other geometries with a valid QC energy are dumped into spares.
2. ADDPOINTS is configured with the desired E_{thresh}, $E_{pot(max)}$, energy window, and other parameters, and started with the aforementioned two geometry sets.
3. The rebuild is complete when ADDPOINTS quits due to not having any more geometries with a SFE > E_{thresh} in spares.

We envision three major use cases when this may be desirable. First, one could use this to prepare the initial fitting set for starting PES development with ROBOSURFER, by generating geometries using traditional methods, running QC computations on all of them, and then performing a rebuild. This could generate both a compact initial fitting set and some spares at the same time.

Second, after one is satisfied with the state of the PES and terminates ROBOSURFER, one could perform a rebuild with the same E_{thresh} and $E_{pot(max)}$, in an attempt to reduce the size of the fitting set. This may be desirable if one ran ROBOSURFER with a

low-cost QC method but wants to switch to a high-accuracy method for the final PES.

Third, one might have developed a PES with high $E_{pot(max)}$ to describe high-energy regions accurately, resulting in a large fitting set. A rebuild performed with a lower $E_{pot(max)}$ could shrink the fitting set considerably, while maintaining accuracy in the lower energy region.

In this work, we present results for the latter two use cases in Section V.

IV. COMPUTATIONAL DETAILS

IV.A. PES Development. We have developed five new full-dimensional analytic PESs for the $CH_3Br + F^-$ system, which we will refer to as PES I, PES IIa, PES IIb, PES IIc, and PES III.

All quantum chemical calculations have been performed using the MOLPRO 2015.1 package.[51−55] The first four PESs use explicitly correlated second-order Møller−Plesset perturbation theory with density fitting[56] (DF-MP2-F12), while PES III uses explicitly correlated coupled cluster singles and doubles with perturbative triples[37] (CCSD(T)-F12b) as the electronic structure theory. All QC computations employ the correlation-consistent polarized valence triple-ζ basis sets of Dunning[57] augmented with diffuse functions,[58] denoted as aug-cc-pVTZ (AVTZ). For bromine, the innermost 10 electrons are replaced by a relativistic effective core potential[59] (denoted as ECP10MDF in MOLPRO), and the corresponding pseudopotential basis set[59] (aug-cc-pVTZ-PP, AVTZ-PP) is used.

To lower the incidence of unphysical energies due to Hartree−Fock misconvergence, the density and energy convergence criteria for the HF iterations were tightened to 10^{-9} and 10^{-10}, respectively. To support this tight convergence, the two-electron integral calculation and storage thresholds were also lowered to 10^{-12}.

By default MOLPRO attempts to freeze orbital occupations at the 20th SCF iteration and does at most 60 iterations. To better handle difficult cases, this was changed to 50 and 100 iterations, respectively.

All PESs have been fitted using the polynomial-invariant-based implementation of the PIP method with Morse-type variables[5] ($y_{ij} = e^{-r_{ij}/a}$), where r_{ij} denotes interatomic distances and a is fixed at 3 bohr. The maximum order of the polynomial expansion was set to 6 in all cases, yielding 10831 coefficients that were determined by a weighted linear least-squares fit. During the fitting, each geometry in the fitting set has a weight given by eq 6:

$$w(E) = \frac{E_{dwt0}}{E + E_{dwt0}} \cdot \frac{E_{dwt1}}{E + E_{dwt1}} \qquad (6)$$

where E is the energy of the geometry relative to the lowest energy in the fitting set and E_{dwt0} and E_{dwt1} are fixed at 0.1 and 0.5 hartree, respectively. E_{dwt0} is also used to define the energy ranges where the RMS fitting errors are given; in this work, we use the $[0, E_{dwt0})$, $[E_{dwt0}, 2E_{dwt0})$, and $[2E_{dwt0}, 5E_{dwt0})$ intervals.

PES I has been developed by running the ROBOSURFER system. The a parameter in the (PI-)EW-RMSD distance metrics (eqs 2 and 3) was set to 2 Å. The energy window had an upper limit of +0.675 hartree (+424 kcal/mol) relative to free reactants and no lower limit. While HOLEBUSTER was enabled, 8 HOLEBUSTER instances were spawned in every ROBOSURFER iteration (Figure 1). QCT trajectories were run at a collision energy of 60 kcal/mol, which is the largest collision energy we set out to accurately model with our PESs. The targeted PES accuracy (E_{thresh}) was

set to 1.5×10^{-3} hartree (0.94 kcal/mol) in all programs, and the maximum energy for full accuracy ($E_{pot(max)}$) was set to +88.6 kcal/mol relative to free reactants, corresponding to the sum of the maximum collision energy, the harmonic zero-point energy (ZPE) of the reactants, and a 5 kcal/mol "safety margin". This attempts to guarantee full accuracy even in the hypothetical scenario where all kinetic energy in the system is transformed into potential energy.

It should be noted that PES I was developed in parallel with the programs of the ROBOSURFER system, and major code changes were typically followed by a PES rebuild. During development, geometries from the fitting set of our previously published PES for the $CH_3I + F^-$ system[60] were mixed into the geometry pool. For a more detailed description please see the SI.

After the fourth such rebuild, the development of the final version of PES I was started from 39073 geometries included in the fit and 166199 spare geometries. For the first 2810 ROBOSURFER iterations both QCT MD and HOLEBUSTER were used to generate new geometries, adding 23444 geometries to the fitting set. At this point, HOLEBUSTER only found holes with a low probability; thus, we disabled running HOLEBUSTER to shift the focus to improving the fitting error of regions most relevant to QCT MD. The development was halted after 2484 further iterations that added 39199 geometries to the fitting set; at this point QCT calculations no longer yielded any impossible products, and the rate at which new points were being added was slowing down.

PESs IIa, IIb, and IIc were all created by merging the fitting set of PES I with the corresponding spares and then rebuilding PES I in different ways.

PES IIa used the same parameters in ADDPOINTS that were used during the development of PES I, and the rebuild was started with a fitting set only containing a single geometry that was randomly selected from the merged set. The observant reader might note that the fit is guaranteed to be underdetermined until the number of fitting points reaches the number of coefficients in the fit. While such a fit is bound to suffer severely from overfitting, there is no technical reason preventing its use in the rebuilding process. PES IIb again used the same parameters, but the single initial geometry was chosen to be the global minimum geometry of the $CH_3Br + F^-$ system (later denoted as POSTMIN).

PES IIc used the same initial geometry as PES IIa, but $E_{pot(max)}$ was decreased to +58.6 kcal/mol, corresponding to a maximal planned collision energy of 30 kcal/mol.

PES III was created by taking the fitting set of PES IIc, computing CCSD(T)-F12b/AVTZ(-PP) energies at those geometries, and running the PES fitter. Some geometries suffered from convergence issues; those were not included in PES III. Two more geometries were manually removed from PES III, due to being extreme outliers.

IV.B. QCT Reaction Dynamics Simulations. We performed QCT computations for the $CH_3Br(v = 0) + F^-$ reaction using all five PESs. The quasiclassical vibrational ground state of CH_3Br was prepared by standard normal mode sampling,[12] and the velocities were adjusted to set the rotational angular momentum to zero.

The initial orientation of the CH_3Br molecule was randomly sampled, and the initial center of mass distance was set to $\sqrt{x^2 + b^2}$, where b is the impact parameter and $x = 40$ bohr. Trajectories were run at collision energies (E_{coll}) of 7.4, 15.9, 35.3, 42.5, 50.0, and 60.0 kcal/mol. The impact parameter was
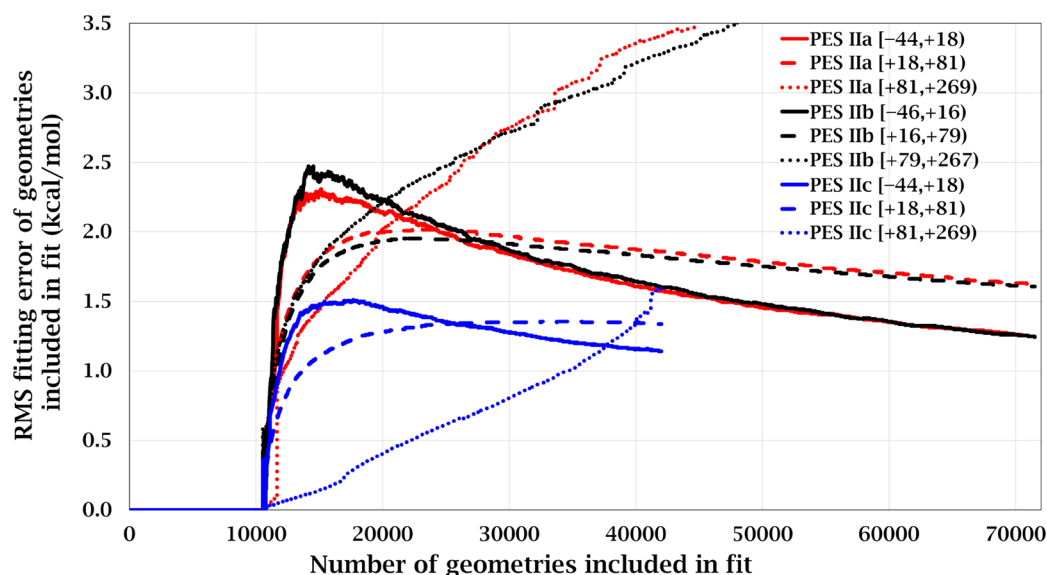
Figure 4. Evolution of RMS fitting errors for geometries in the fitting set, during the rebuilding processes that yielded PESs IIa, IIb, and IIc. All energies are in kcal/mol, and the energy range end points are relative to the energy of free reactants.

scanned from 0 bohr to b_{max} with a step size of 0.5 bohr, and the maximum impact parameter at which reaction still occurs (b_{max}) was determined for every PES and every collision energy individually. The values of b_{max} can be seen in Table S1.

For each PES, for every combination of b and E_{coll} 5000 trajectories have been run, meaning roughly 3.25 million trajectories in total. The trajectories have been propagated using the velocity Verlet integrator,[61] with a time step of 3 atomic units (0.0726 fs). Every trajectory has been propagated until the largest interatomic distance became 1 bohr larger than the largest initial one. We used the Avogadro molecule editor[62,63] for visualizing stationary points and trajectories.

The trajectories were analyzed with a new in-house tool that can detect unphysical products and product geometries that cannot be clearly assigned to any product channel. The integral cross sections (ICSs) were obtained by a b-weighted numerical integration of the reaction probabilities over b. The integration was performed with the trapezoidal rule.

V. RESULTS AND DISCUSSION

V.A. PES Development Using ROBOSURFER. The fitting set of PES I contains 101716 geometries, while the corresponding spares consists of 242074 geometries. This fitting set is roughly twice as large as the fitting sets we have used in previous studies for similar systems.[4,60] The number of spares is also substantial; we attribute this to a lack of filtering in early development versions of ROBOSURFER. While one may consider spares to be wasted CPU cycles, it should be noted that the fact of them being spares indicates that there are 242074 geometries outside the fitting set with a scaled fitting error less than 0.94 kcal/mol. This suggests that PES I is highly accurate even at geometries outside the fitting set.

The RMS fitting error for the fitting set of PES I is 1.04 kcal/mol for geometries between −45.2 and +17.6 kcal/mol (relative to free reactants), 1.46 kcal/mol between +17.6 and +80.3, and 5.4 kcal/mol between +80.3 and +268.6 kcal/mol. The observant reader might note that these RMS fitting errors are larger than what is typically reported for reactive PIP PESs, despite the high order of the polynomial fit. We believe this is due to the combination of three factors.

First, previous PESs have not been crafted to only contain geometries that have a high fitting error. The selective addition process done by ADDPOINTS means that geometries that have a low fitting error and would otherwise artificially lower the RMS error statistic are mostly excluded from the fit.

Second, previous PESs may have undergone less sampling effort, especially in high energy regions, while PES I was extensively sampled and extra effort was put into fixing holes. It is plausible that trying to fit a PES that is correct over the entire configuration space (even if only qualitatively correct) while keeping RMS errors low is a very challenging task for PES fitting methods.

Third, we suspect that some of the excess fitting error may be attributable to weaknesses of the underlying electronic structure methods. Over the course of the development of PES I, over 6833 geometries were discarded due to Hartee–Fock non-convergence or a failed sanity check in MOLPRO's MP2 module. This clearly indicates that ROBOSURFER ventured into regions of the configuration space where standard single-reference methods start to become less reliable. Thus, it is conceivable that, despite our efforts to avoid this, some of the geometries in the fitting set and spares may have energies compromised by multireference effects or HF misconvergence. The number of CCSD-F12 convergence failures that happened when the fitting set of PES IIc was recalculated reinforces these suspicions. This could also explain the large number of fitting points required, as more geometries may have been added by the ROBOSURFER system to compensate for fitting points with spurious energies.

V.B. PESs from Rebuilding. PES IIa has a fitting set of 71158 geometries, 30558 less than PES I (a 30% reduction), while the corresponding set of spares is larger by the same amount, since no geometries were added or discarded. RMS fitting errors are 1.25 kcal/mol between −44.4 and +18.4 kcal/mol, 1.63 kcal/mol from +18.4 to +81.1, and 4.46 kcal/mol between +81.1 and +269.4 kcal/mol. Compared to PES I, RMS errors increased in the lower two energy ranges, while they decreased in the top range.

PES IIb has a fitting set of 71508 geometries, marginally smaller than that of PES IIa, suggesting that the choice of initial geometry is noncritical for the success of the rebuilding process.
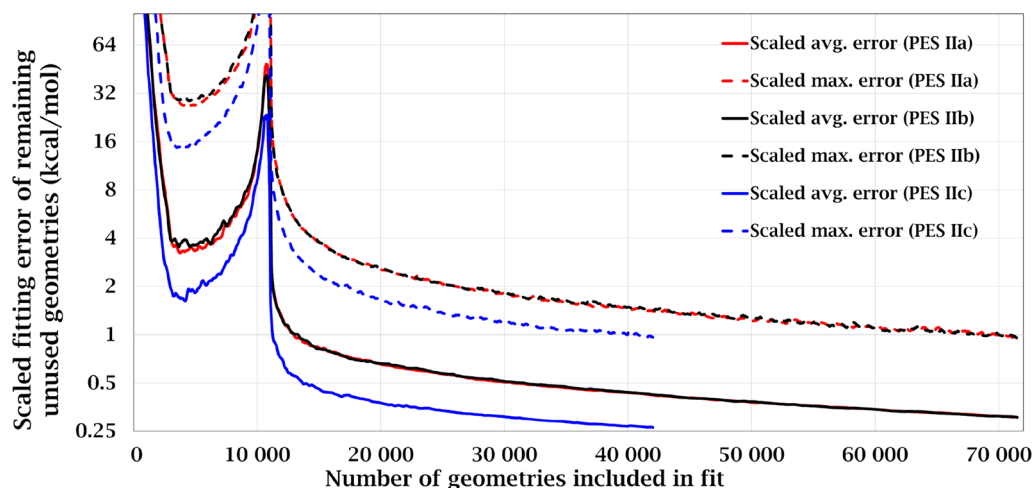
Figure 5. Evolution of scaled average and maximum fitting errors for geometries not yet included in the fitting set, during the rebuilding processes that yielded PESs IIa, IIb, and IIc. The scaling function is eq 5, and average and maximum errors were smoothed with 150 point moving averages. Note the logarithmic scale.
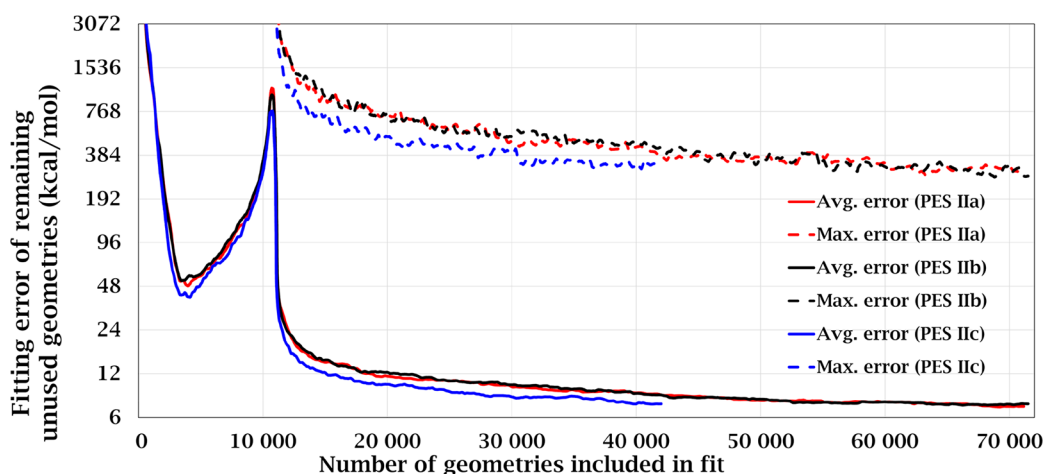
Figure 6. Evolution of average and maximum fitting errors for geometries not yet included in the fitting set, during the rebuilding processes that yielded PESs IIa, IIb, and IIc. Average and maximum errors were smoothed with 150 point moving averages. Note the logarithmic scale.

The RMS errors, 1.24 kcal/mol between −46.3 and +16.5 kcal/mol, 1.61 kcal/mol from +16.5 to +79.2, and 4.34 kcal/mol between +79.2 and +267.5 kcal/mol, confirm this, as the fitting errors are only slightly lower than in the case of PES IIa. This may be attributable to a change in the weights (eq 6) of every geometry, caused by the lower minimum energy in the fit (since PES IIb was rebuilt starting with the global minimum geometry). The shift in the range boundaries is also due to the change in minimum energy.

PES IIc has only 42012 fitting points (a 41% reduction compared to PES IIa), showing that one can greatly reduce the number of fitting points used if accuracy is only required in a more modest energy range. The RMS errors are also lower, 1.14 kcal/mol between −44.4 and +18.4 kcal/mol, 1.34 kcal/mol from +18.4 to +81.1, and 1.59 kcal/mol between +81.1 and +269.4 kcal/mol, suggesting it is much more challenging to achieve low fitting errors over a wider energy range. The reduction in RMS error in the upper two energy ranges is partially attributable to having fewer geometries in those regions.

The progress of the rebuilding processes that yielded PESs IIa, IIb, and IIc can be seen in Figure 4 and Figure 5.

Below 10831 geometries, there are more polynomial coefficients than data points, the fit is underdetermined, the fitted PES passes through all fitting points exactly, and thus their RMS error is zero within numerical error (Figure 4). This RMS error statistic is of course misleading, as for geometries outside the fitting set the fitting error is much larger due to overfitting.

After the fit ceases to be underdetermined, RMS errors in all energy ranges rise sharply, and typically for a while the lowest energy region has the highest RMS error, while the highest energy region has the lowest RMS error in the fitting set. This is unusual, as the weighting used in the PES fitter program (eq 6) strongly biases the fit to be more accurate at lower energies. We attribute this to the interplay between the evolution of the fit and the geometry addition strategy of ADDPOINTS.

After the initial sharp rise, the RMS errors in the lower energy regions pass through a maximum and begin to fall, crossing the RMS error of the middle energy region around 26000 points. The RMS error continues to rise in the highest energy region, while in the middle region it slowly decreases (PESs IIa and IIb) or plateaus (PES IIc). It is interesting to note that the RMS errors in the lowest region kept decreasing until the end of the rebuild, suggesting that lower RMS errors could have been
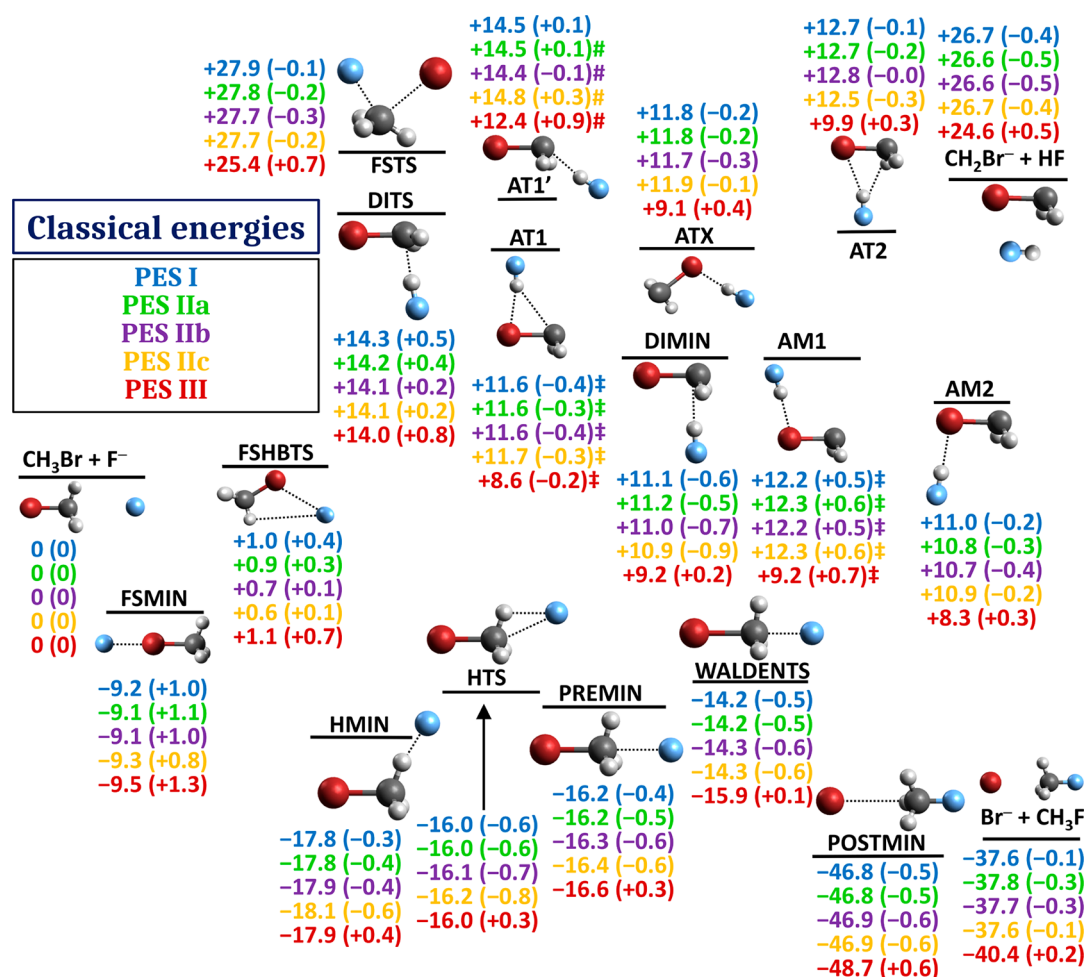
Figure 7. Relative classical energies of products and stationary points of the $CH_3Br + F^-$ system using the five different PESs. Fitting errors are given in parentheses as the difference in optimized energies between a PES and its parent QC method. All energies are in kcal/mol. ‡: PES energy at QC geometry. #: The optimization converged to a second-order saddle point.

achieved (in regions of chemical interest) if spares contained more geometries that could have been added, or if the target accuracy (E_{thresh}) was set to a lower value.

Looking at the fitting errors of the fitting set gives only half the picture, as geometries not included may have a large error even if the RMS errors of the fitting set are low. Figure 5 shows that, after leaving underdeterminedness behind, the SFEs of the remaining spares decrease in a gradually slowing manner.

PES IIa and PES IIb show negligible differences in SFEs throughout the rebuilds, supporting that the choice of initial geometry is noncritical. The SFEs of PES IIc are consistently lower, suggesting that lowering $E_{pot(max)}$ can result in a PES that is more accurate, albeit only over a more limited energy range.

Figure 6 shows that the unscaled errors exhibit a very similar evolution over the course of the rebuild. PES IIa and PES IIb are once again nearly indistinguishable, while PES IIc provides lower errors for fitting sets of the same size. The latter is noteworthy, because these unscaled errors are only affected by $E_{pot(max)}$ indirectly, by influencing which geometries are included in the fit. This suggests that including a lower percentage of high energy geometries paradoxically results in a better description of high energy regions. It is also interesting to note that PES IIc reached the same average unscaled error (7.4 kcal/mol) as PES IIb, despite the former using ~40% less geometries in the fitting set.

V.C. PES III. Out of the 42012 geometries used in PES IIc, 646 had to be discarded due to coupled cluster convergence issues; thus, PES III originally would have had 41366 fitting points. This PES, however, had unexpectedly large fitting errors at known stationary points, prompting us to manually remove the two most extreme outliers (Figure S1). These geometries also had large T1 diagnostic[64] values, suggesting a breakdown of single-reference coupled cluster theory.

The removal of just these two points reduced RMS fitting errors substantially in the top energy range but had almost no effect on the lower two ranges. The RMS fitting errors of the fitting set changed from 0.84 to 0.83 between −47.5 and +15.3 kcal/mol, from 1.08 to 1.06 kcal/mol between +15.3 and +78.0 kcal/mol, and from 0.97 to 0.38 kcal/mol between +78.0 and +266.3 kcal/mol. Despite the small change in the RMS error of the lower ranges, the fitting errors of stationary points have also improved by a couple tenths of a kcal/mol upon the removal of the two outliers.

The observant reader might note that the RMS fitting errors of PES III are significantly lower than the corresponding values of PES IIc or indeed any of the other PESs presented in this work. We attribute this primarily to the removal of the 646+2 geometries from the fitting set, where the coupled cluster iterations could not easily converge to a reasonable solution. We expect that at any such geometry the DF-MP2-F12 energy is also
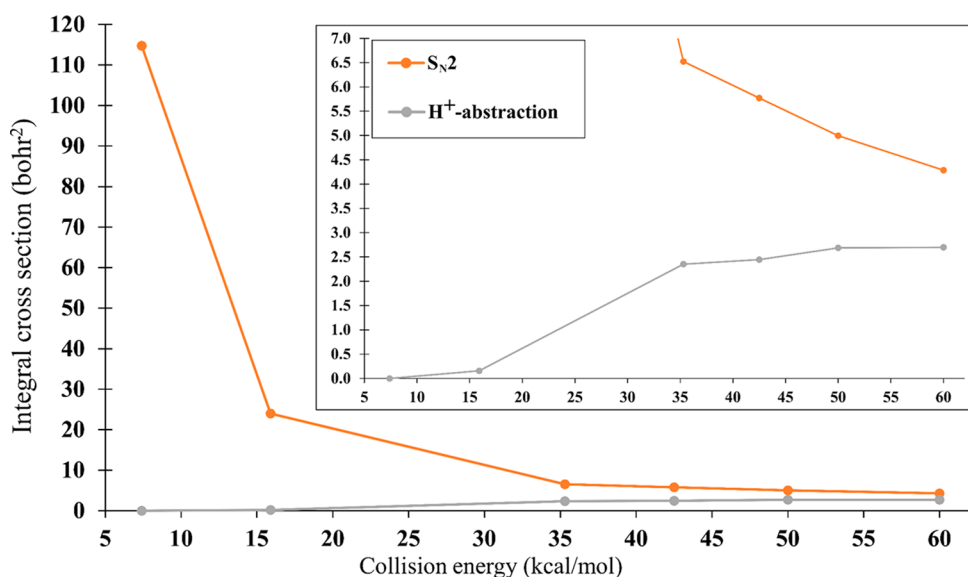
Figure 8. Integral cross sections for the S_N2 and proton-abstraction product channels of the $CH_3Br(v = 0) + F^-$ reaction, at different collision energies, using PES I.

likely to be inaccurate, and the inclusion of such geometries could be partially responsible for the fitting error of the other 4 PESs.

While coupled cluster nonconvergence may have removed some of the most problematic geometries, the breakdown of traditional single reference coupled cluster methods at far-from-equilibrium structures is well-known,[38,65] and it is possible that PES III still contains geometries with spurious energies in its fitting set. Testing methods for expunging such geometries are however beyond the scope of this work.

V.D. Accuracy of the PESs at Known Stationary Points. Figure 7 depicts all known stationary points of the $CH_3Br + F^-$ system (most of them determined in the present study for the first time) and the two lowest energy product channels (S_N2 and proton abstraction). Fitting errors at these geometries are generally excellent for all PESs and stay below 1 kcal/mol at all geometries except the halogen-bonded FSMIN, even though no effort was made to enhance the sampling of known stationary points.

Two geometries in the H-abstraction region, AT1 and AM1, could not be located on any of the PESs. This region is crowded with critical points that are very close in potential energy and even the small fitting errors of the PESs are enough to make them vanish. For this reason, the energies and fitting errors reported at these two points have been computed using geometries optimized with the parent QC methods of the PESs, instead of optimizing them using the PESs.

Another geometry, AT1′, is slightly problematic on all PESs except PES I: while the fitting errors are in the expected range, the optimized geometries have two imaginary normal modes. It is worth noting that this curvature artifact is present on PESs IIa/IIb, despite the excellent fitting error of ±0.1 kcal/mol at this geometry. This strange curvature seems to have no significant effect on ZPE corrected relative energies (see Figure S2).

For minima and the two product channels, PES I achieved an RMS error of 0.5 and 0.4 kcal/mol for saddle points (SPs). PESs IIb and IIc both achieved 0.6 and 0.4 kcal/mol for the same statistics, while PES IIa did slightly better at SPs with an RMS error of 0.3 kcal/mol. These errors are significantly better than the RMS fitting error of the fitting sets of these PESs and are also

below the targeted PES accuracy (0.94 kcal/mol) used during the PES development and rebuilds.

PES III achieved an RMS error of 0.6 kcal/mol for minima and products and 0.5 kcal/mol for saddle points, slightly worse than PES IIc, despite having a significantly lower RMS error over the fitting set. This suggests that the overall quality of the fitted PES and the RMS errors over the fitting set may have a less-than-perfect correlation.

The ZPE corrected relative energies paint a similar picture; fitting errors are usually low, with the exception of FSMIN (see Figure S2).

V.E. QCT Dynamics Results. The quasiclassical trajectory (QCT) dynamics results from PES I are generally in line with the expectations based on results for the $CH_3Cl + F^{-[4]}$ and $CH_3I + F^{-[60,66]}$ systems. As it is typical for S_N2 reactions with a submerged Walden-inversion barrier, the S_N2 cross section is initially very large but decreases steeply with increasing collision energy (Figure 8).

The proton-abstraction channel starts to open slightly at a collision energy of 15.9 kcal/mol; this is well below the adiabatic reaction energy of +23.4 kcal/mol found on PES I, and therefore these reactions must be ZPE violating. This is in agreement with the other two systems, where small but nonzero ZPE violating abstraction cross sections were found. At 35.3 kcal/mol E_{coll} the abstraction channel is fully open, and further increases in E_{coll} result in only modest enhancement of the abstraction cross section.

Taking a look at the QCT results from the other 4 PESs (Figure 9), the overall theme is that all PESs using DF-MP2-F12 as their parent QC method perform similarly, with sporadic differences, while PES III which uses CCSD(T)-F12b often yields drastically different cross sections. The latter is in line with our previous findings,[13] where we also saw a marked enhancement of both the S_N2 and the abstraction cross sections for the $CH_3I + F^-$ system.

On a closer inspection, the S_N2 cross sections for PESs IIa/IIb/IIc usually stay within ±5% of the PES I results (Figure S3), with the exception of the anomalous results at 42.5 kcal/mol and the drop in ICS for PES IIc at 60 kcal/mol. For the H^+-abstraction channel the differences tend to be larger (Figure S4).
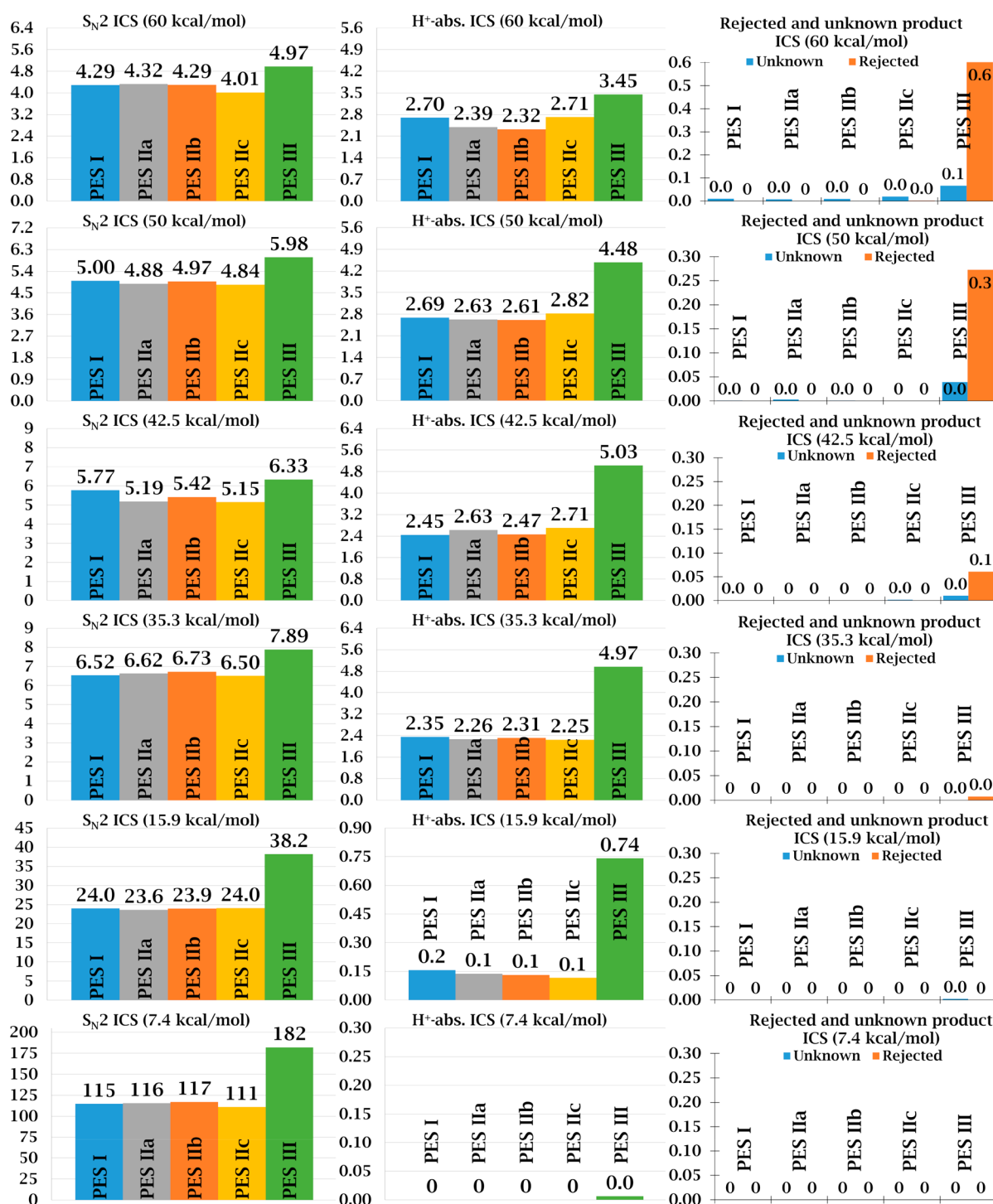
Figure 9. Integral cross sections for the S_N2, proton abstraction, rejected, and unknown product channels of the $CH_3Br(v = 0) + F^-$ reaction, at six different collision energies, using the five PESs developed. All cross sections are in bohr^2.

There are two more pseudochannels shown in Figure 9 that are important to discuss, unknown products and rejected products. Unknown products could not be unambiguously assigned to any known product channel; these are typically vibrationally highly excited minor products, such as a CH_2Br^- product with a C−Br stretching vibration excited almost to the point of dissociation. The cross section of the unknown channel is generally negligible for all PESs and collision energies and only exceeds 0.05 bohr^2 for PES III at 60 kcal/mol E_{coll}.

With that said, the slight increase in unknown and rejected trajectories for PES IIc at 60 kcal/mol E_{coll}, together with the S_N2 ICS of PES IIc being somewhat of an outlier at the same energy, suggests that PES IIc starts to become less reliable beyond 50 kcal/mol E_{coll}. This is a better-than-expected result, as PES IIc was rebuilt with a maximum E_{coll} target of 30 kcal/mol, suggesting that the heuristic we used to estimate the full accuracy limit ($E_{pot(max)}$ = 58.6 kcal/mol) might be a significant overestimation. If that is indeed the case, the size of the fitting set

could further be reduced, while maintaining the desired accuracy up to 30 kcal/mol of E_{coll}.

Rejected products are thermodynamically impossible products that significantly violate energy conservation, such trajectories are a typical result of having holes on a PES, and thus the ICS of this channel is an indicator of PES quality. For most PESs, this channel has an entirely negligible cross section at all energies, but for PES III this is not true.

From 60 to 42.5 kcal/mol E_{coll}, PES III has a significant rejected cross section and impossible products are still found at 35.3 kcal/mol. This clearly shows that recalculating the fitting set of PES IIc with CCSD(T)-F12b introduced additional artifacts, despite having a better RMS error over its fitting set. It remains to be seen if this effect is due to the presence of geometries in the fitting set where the single-reference (T) approximation breaks down or if this is a genuine limitation of constructing a high-accuracy PES by simply recalculating a known good PES with a higher level of electronic structure theory. If the latter case is true, it might be necessary to run a finishing pass of ROBOSURFER with the desired high-accuracy method.

VI. SUMMARY, CONCLUSIONS, AND OUTLOOK

The construction of reliable, highly accurate, global potential energy surfaces from *ab initio* data has been a major challenge for decades. Advances in computer hardware, electronic structure theory, and PES fitting methods have greatly alleviated many challenges in PES construction, but the generation of suitable fitting sets has remained a bottleneck so far.

We have developed the initial version of the ROBOSURFER program system that completely automates the generation of fitting points, performs *ab initio* computations, and iteratively improves the PES under development. The program system is highly modular and can be easily adapted to accommodate future advances in PES fitting, as it does not rely on the fitting method to supply uncertainties.

ROBOSURFER does not require any significant assumptions or knowledge of the stationary points or the reaction paths of the system; indeed it is not even necessary to know all possible product channels in advance of PES development.

ROBOSURFER also includes techniques for minimizing the number of *ab initio* computations required, avoiding geometries with a history of convergence issues, fixing unphysical minima (holes) on the PES, and minimizing the size of the fitting set.

As an offshoot of ROBOSURFER we have also devised a rebuilding scheme that can effectively reduce the size of a fitting set.

We have demonstrated the efficacy of automated PES development, by developing a PES for the $CH_3Br + F^-$ system at the DF-MP2-F12/aug-cc-pVTZ level, with a target accuracy of 0.94 kcal/mol and a desired maximum collision energy of 60 kcal/mol.

This PES (PES I) performed excellently, the fitting errors of known stationary points are typically well below 1 kcal/mol, and QCT dynamics results obtained with this PES are in line with the results expected for such a system. Of special note is the lack of impossible products, even at 60 kcal/mol of collision energy, suggesting a successful elimination of holes on the PES.

We have also tested our rebuilding scheme, by creating three more PESs, each rebuilt differently. PES IIa and PES IIb have been rebuilt using the same parameters but a different starting geometry. PES IIa has been rebuilt from a geometry randomly drawn from the pool of geometries used for the rebuild, whereas

PES IIb used the global minimum structure of the system. Despite this difference, PESs IIa/IIb both use 30% smaller fitting sets, have very similar fitting errors, and yield similar QCT results.

PES IIc used the same initial geometry as IIa but only 30 kcal/mol of desired collision energy. This PES uses a fitting set 59% smaller than PES I, yet it retains almost all of its accuracy at stationary points and appears to be suitable for QCT calculations with collision energies up to 50 kcal/mol.

We have also taken the fitting set of PES IIc and recomputed it at the CCSD(T)-F12b/aug-cc-pVTZ(-PP) level for our fifth PES in this work, PES III. While out of all five PESs this PES has the lowest RMS fitting errors over its fitting set, it also has an RMS error at stationary points slightly higher than that of PES IIc. QCT results are also mixed: on the one hand they are in line with our previous results,[13] showing an enhancement of both S_N2 and proton-abstraction cross sections, but on the other hand we see impossible products at collision energies as low as 35.3 kcal/mol, indicating the presence of holes.

We have found that some geometries from the fitting set of PES IIc failed to converge to a reasonable coupled cluster solution. It is reasonable to assume that at geometries where the perturbative (T) approximation gives poor results the corresponding MP2 energy is also of questionable quality.

We believe the holes that seem to be present in PES III could potentially be remedied by using an electronic structure method more robust than CCSD(T). In a previous study[13] we have seen some limited evidence that novel single-reference coupled cluster theories which handle the onset of multireference effects more gracefully (such as quasi-variational coupled-cluster theory[38,67]) may be useful for overcoming such problems. This is a direction we plan to investigate in detail, as we have run into the disruptive effects of spurious (T) energies on multiple occasions.

Overall, we make the following conclusions:

(a) The ROBOSURFER program system described in this work is effective at automatically developing nontrivial reactive PESs.

(b) The rebuilding scheme proposed in this work can yield compact fitting sets without sacrificing accuracy for compactness, provided the pool of *ab initio* data used for the rebuild is sufficiently large and diverse.

(c) Such pools of data can be generated by running ROBOSURFER.

(d) The rebuild process is insensitive to the seed geometry, at least in the case of a single seed geometry.

(e) Rebuilding with a lower desired collision energy can greatly reduce the size of the fitting set, while maintaining most of the accuracy.

(f) Recalculating the fitting set of a PES that has been developed at a lower level is an expedient way for creating a PES that uses a high-accuracy *ab initio* method, but PESs created in such a manner may suffer from artifacts.

(g) It remains to be seen if such artifacts are avoidable or running a finishing pass of ROBOSURFER is necessary.

(h) Geometries where traditional single reference methods break down pose a serious hazard to the accuracy of fitted PESs.

The present study shows that the fully automated development and testing of substantially complicated reactive PESs is within reach. The current version of ROBOSURFER is already being successfully used in our group for the development of multiple

new PESs for systems with 7−9 atoms, such as $CH_3I + OH^-$, $C_2H_6 + F^•/Cl^•$, and $C_2H_5Cl + F^-$.

It is not too hard to envision a program system that monitors the fitting errors of all known stationary points, minimum energy paths, and product channels, as well as QCT results, stopping PES development when these values converge.

If problems related to the unphysical behavior of single-reference methods can be overcome (or circumvented), then such a program system could potentially operate in a nearly black-box fashion, making high-quality *ab initio* PES development much more accessible for the computational chemistry community. We plan to extend the ROBOSURFER system toward this direction and make all source code publicly available in the future.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.9b01006.

> Additional technical details of the ROBOSURFER program system and its components (PDF)
>
> Table of acronyms and additional figures and tables (PDF)
>
> Potential energies and harmonic ZPEs obtained by MOLPRO and the five fitted PESs (PDF)
>
> Optimized geometries obained by MOLPRO and the five fitted PESs (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: tibor.gyori@chem.u-szeged.hu (T.G.).
*E-mail: gczako@chem.u-szeged.hu (G.C.).

ORCID Ⓘ

Tibor Győri: 0000-0002-4078-7624
Gábor Czakó: 0000-0001-5136-4777

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Dóra Papp for reviewing the manuscript before submission.

■ REFERENCES

(1) McNaught, A. D.; Wilkinson, A. Collision theory. In *IUPAC. Compendium of chemical terminology*, 2nd ed.; Chalk, S. J., Ed.; Online version (2019−); Blackwell Scientific Publications: Oxford, U.K., 1997; DOI: 10.1351/goldbook.C01170.

(2) Mikosch, J.; Trippel, S.; Eichhorn, C.; Otto, R.; Lourderaj, U.; Zhang, J. X.; Hase, W. L.; Weidemüller, M.; Wester, R. Imaging nucleophilic substitution dynamics. *Science* **2008**, *319*, 183−186.

(3) Ingold, C. K. *Structure and Mechanisms in Organic Chemistry*; Cornell University Press: Ithaca, NY, 1953.

(4) Szabó, I.; Czakó, G. Revealing a double-inversion mechanism for the $F^- + CH_3Cl$ S_N2 reaction. *Nat. Commun.* **2015**, *6*, 5972.

(5) Braams, B. J.; Bowman, J. M. Permutationally invariant potential energy surfaces in high dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577−606.

(6) Collins, M. A. Molecular potential-energy surfaces for chemical reaction dynamics. *Theor. Chem. Acc.* **2002**, *108*, 313−324.

(7) Dawes, R.; Passalacqua, A.; Wagner, A. F.; Sewell, T. D.; Minkoff, M.; Thompson, D. L. Interpolating moving least-squares methods for fitting potential energy surfaces: Using classical trajectories to explore configuration space. *J. Chem. Phys.* **2009**, *130*, 144107.

(8) Chen, J.; Xu, X.; Xu, X.; Zhang, D. H. A global potential energy surface for the $H_2 + OH \leftrightarrow H_2O + H$ reaction using neural networks. *J. Chem. Phys.* **2013**, *138*, 154301.

(9) Li, J.; Jiang, B.; Guo, H. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. II. Four-atom systems. *J. Chem. Phys.* **2013**, *139*, 204103.

(10) Patrício, M.; Santos, J. L.; Patrício, F.; Varandas, A. J. C. Roadmap to spline-fitting potentials in high dimensions. *J. Math. Chem.* **2013**, *51*, 1729−1746.

(11) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457−484.

(12) Hase, W. L. Classical trajectory simulations: Final conditions, classical trajectory simulations: Initial conditions. In *Encyclopedia of computational chemistry*, 1st ed.; von Rague Schleyer, P.; John Wiley & Sons Ltd.: Chichester, U.K., 1998; Vol. *1*, pp 399−407.

(13) Győri, T.; Olasz, B.; Paragi, G.; Czakó, G. Effects of the level of electronic structure theory on the dynamics of the $F^- + CH_3I$ reaction. *J. Phys. Chem. A* **2018**, *122*, 3353−3364.

(14) Ischtwan, J.; Collins, M. A. Molecular potential energy surfaces by interpolation. *J. Chem. Phys.* **1994**, *100*, 8080−8088.

(15) Dawes, R.; Wagner, A. F.; Thompson, D. L. *Ab initio* wavenumber accurate spectroscopy: $^1CH_2$ and HCN vibrational levels on automatically generated IMLS potential energy surfaces. *J. Phys. Chem. A* **2009**, *113*, 4709−4721.

(16) Majumder, M.; Ndengue, S. A.; Dawes, R. Automated construction of potential energy surfaces. *Mol. Phys.* **2016**, *114*, 1−18.

(17) Xie, Z.; Bowman, J. M. Permutationally invariant polynomial basis for molecular energy surface fitting via monomial symmetrization. *J. Chem. Theory Comput.* **2010**, *6*, 26−34.

(18) Bowman, J. M.; Czakó, G.; Fu, B. High-dimensional *ab initio* potential energy surfaces for reaction dynamics calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 8094−8111.

(19) Qu, C.; Yu, Q.; Bowman, J. M. Permutationally invariant potential energy surfaces. *Annu. Rev. Phys. Chem.* **2018**, *69*, 151−175.

(20) Guan, Y.; Yang, S.; Zhang, D. H. Construction of reactive potential energy surfaces with Gaussian process regression: active data selection. *Mol. Phys.* **2018**, *116*, 823−834.

(21) Cui, J.; Krems, R. V. Efficient non-parametric fitting of potential energy surfaces for polyatomic molecules with Gaussian processes. *J. Phys. B: At., Mol. Opt. Phys.* **2016**, *49*, 224001.

(22) Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.

(23) Manzhos, S.; Dawes, R.; Carrington, T. Neural network-based approaches for building high dimensional and quantum dynamics-friendly potential energy surfaces. *Int. J. Quantum Chem.* **2015**, *115*, 1012−1020.

(24) Handley, C. M.; Popelier, P. L. A. Potential energy surfaces fitted by artificial neural networks. *J. Phys. Chem. A* **2010**, *114*, 3371−3383.

(25) Qu, C.; Yu, Q.; Van Hoozen, B. L.; Bowman, J. M.; Vargas-Hernández, R. A. Assessing Gaussian process regression and permutationally invariant polynomial approaches to represent high-dimensional potential energy surfaces. *J. Chem. Theory Comput.* **2018**, *14*, 3381−3396.

(26) Jiang, B.; Guo, H. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. *J. Chem. Phys.* **2013**, *139*, 054112.

(27) Shao, K.; Chen, J.; Zhao, Z.; Zhang, D. H. Communication: Fitting potential energy surfaces with fundamental invariant neural network. *J. Chem. Phys.* **2016**, *145*, 071101.

(28) Czakó, G.; Bowman, J. M. Dynamics of the reaction of methane with chlorine atom on an accurate potential energy surface. *Science* **2011**, *334*, 343−346.

(29) Frankcombe, T. J.; Collins, M. A.; Worth, G. A. Converged quantum dynamics with modified Shepard interpolation and Gaussian wave packets. *Chem. Phys. Lett.* **2010**, *489*, 242−247.

(30) Szabó, I.; Császár, A. G.; Czakó, G. Dynamics of the F⁻ + CH₃Cl → Cl⁻ + CH₃F S$_N$2 reaction on a chemically accurate potential energy surface. *Chem. Sci.* **2013**, *4*, 4362−4370.

(31) Wang, Y.; Song, H.; Szabó, I.; Czakó, G.; Guo, H.; Yang, M. Mode-specific S$_N$2 reaction dynamics. *J. Phys. Chem. Lett.* **2016**, *7*, 3322−3327.

(32) Zhang, W.; Zhou, Y.; Wu, G.; Lu, Y.; Pan, H.; Fu, B.; Shuai, Q.; Liu, L.; Liu, S.; Zhang, L.; Jiang, B.; Dai, D.; Lee, S.-Y.; Xie, Z.; Braams, B. J.; Bowman, J. M.; Collins, M. A.; Zhang, D. H.; Yang, X. Depression of reactivity by the collision energy in the single barrier H + CD₄ -> HD + CD₃ reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 12782−12785.

(33) Li, J.; Jiang, B.; Song, H.; Ma, J.; Zhao, B.; Dawes, R.; Guo, H. From *ab initio* potential energy surfaces to state-resolved reactivities: X + H₂O ↔ HX + OH [X = F, Cl, and O(³P)] reactions. *J. Phys. Chem. A* **2015**, *119*, 4667−4687.

(34) Welsch, R.; Manthe, U. Full-dimensional and reduced-dimensional calculations of initial state-selected reaction probabilities studying the H + CH₄ → H₂ + CH₃ reaction on a neural network PES. *J. Chem. Phys.* **2015**, *142*, 064309.

(35) Li, Y.; Wang, Y.; Wang, D. Quantum dynamics study of the potential energy minima effect on energy efficiency for the F⁻ + CH₃Cl → FCH₃ + Cl⁻ reaction. *J. Phys. Chem. A* **2017**, *121*, 2773−2779.

(36) Sun, L.; Song, K.; Hase, W. L. A S$_N$2 reaction that avoids its deep potential energy minimum. *Science* **2002**, *296*, 875−878.

(37) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(38) Robinson, J. B.; Knowles, P. J. Breaking multiple covalent bonds with Hartree−Fock-based quantum chemistry: Quasi-variational coupled cluster theory with perturbative treatment of triple excitations. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6729−6732.

(39) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.

(40) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.

(41) Quintas-Sánchez, E.; Dawes, R. AUTOSURF: A freely available program to construct potential energy surfaces. *J. Chem. Inf. Model.* **2019**, *59*, 262−271.

(42) Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angew. Chem., Int. Ed.* **2017**, *56*, 12828−12840.

(43) Abbott, A. S.; Turney, J. M.; Zhang, B.; Smith, D. G. A.; Altarawy, D.; Schaefer, H. F. PES-Learn: An open-source software package for the automated generation of machine learning models of molecular potential energy surfaces. *J. Chem. Theory Comput.* **2019**, *15*, 4386−4398.

(44) Qu, C.; Bowman, J. M. An *ab initio* potential energy surface for the formic acid dimer: zero-point energy, selected anharmonic fundamental energies, and ground-state tunneling splitting calculated in relaxed 1−4-mode subspaces. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24835−24840.

(45) Szabó, I.; Czakó, G. Dynamics and novel mechanisms of S$_N$2 reactions on *ab initio* analytical potential energy surfaces. *J. Phys. Chem. A* **2017**, *121*, 9005−9019.

(46) Xie, Z.; Qu, C.; Wang, Q. GitHub repository of MSA-2.0. https://github.com/szquchen/MSA-2.0 (accessed December 10, 2019); latest git commit hash: 0f5ae29b77de000a110836601-b0748eaf212bf8e.

(47) Schryber, J. H.; Polyansky, O. L.; Jensen, P.; Tennyson, J. On the spectroscopically determined potential energy surfaces for the electronic ground states of NO₂ and H₂O. *J. Mol. Spectrosc.* **1997**, *185*, 234−243.

(48) Metz, M. P.; Piszczatowski, K.; Szalewicz, K. Automatic generation of intermolecular potential energy surfaces. *J. Chem. Theory Comput.* **2016**, *12*, 5895−5919.

(49) Nandi, A.; Qu, C.; Bowman, J. M. Using gradients in permutationally invariant polynomial potential fitting: A demonstration for CH₄ using as few as 100 configurations. *J. Chem. Theory Comput.* **2019**, *15*, 2826−2835.

(50) Heap, B. R. Permutations by interchanges. *Comput. J.* **1963**, *6*, 293−298.

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

(52) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. *MOLPRO*, version 2015.1, a package of *ab initio* programs; TTI GmbH: Stuttgart, Germany, 2015; https://www.molpro.net/.

(53) Lindh, R.; Ryu, U.; Liu, B. The reduced multiplication scheme of the Rys quadrature and new recurrence relations for auxiliary function based two-electron integral evaluation. *J. Chem. Phys.* **1991**, *95*, 5889−5897.

(54) Eckert, F.; Pulay, P.; Werner, H.-J. *Ab initio* geometry optimization for large molecules. *J. Comput. Chem.* **1997**, *18*, 1473−1483.

(55) Sun, J.-Q.; Ruedenberg, K. Locating transition states by quadratic image gradient descent on potential energy surfaces. *J. Chem. Phys.* **1994**, *101*, 2157−2167.

(56) Werner, H.-J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, 164102.

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

(58) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(59) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16−18 elements. *J. Chem. Phys.* **2003**, *119*, 11113−11123.

(60) Olasz, B.; Szabó, I.; Czakó, G. High-level *ab initio* potential energy surface and dynamics of the F⁻ + CH₃I S$_N$2 and proton-transfer reactions. *Chem. Sci.* **2017**, *8*, 3164−3170.

(61) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637−649.

(62) *Avogadro*, version 1.20n; 2016; source code and binaires are available at http://avogadro.cc/ (accessed December 10, 2019).

(63) Hanwell, M. D; Curtis, D. E; Lonie, D. C; Vandermeersch, T.; Zurek, E.; Hutchison, G. R Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminf.* **2012**, *4*, 17.

(64) Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36*, 199−207.

(65) Taube, A. G. Alternative perturbation theories for triple excitations in coupled-cluster theory. *Mol. Phys.* **2010**, *108*, 2951−2960.

(66) Mikosch, J.; Zhang, J.; Trippel, S.; Eichhorn, Ch.; Otto, R.; Sun, R.; de Jong, W. A.; Weidemüller, M.; Hase, W. L.; Wester, R. Indirect

dynamics in a highly exoergic substitution reaction. *J. Am. Chem. Soc.* **2013**, *135*, 4250−4259.

(67) Black, J. A.; Knowles, P. J. Quasi-variational coupled-cluster theory: Performance of perturbative treatments of connected triple excitations. *J. Chem. Phys.* **2018**, *148*, 194102.