Methodology for the modeling of chemical reactions on accurate potential energy surfaces

Michael R. Salazar

Department of Chemistry
Union University
1050 Union University Dr.
Jackson, TN 38343
msalazar@uu.edu

Introduction

The development of methods for the accurate description of how a chemical system changes in time is a very important field, but one that seems to have developed more slowly than methods for describing the stationary state properties of systems. This is very unfortunate since being able to accurately describe, from first principles (*ab initio*), collisional processes and the states of any products formed qualifies as a necessary step for a complete understanding of the chemical sciences.

One of the primary difficulties in being able to calculate the dynamics of a chemical system involves the accurate representation of the potential energy surface (PES) for the system.[1] The PES may be rather simple for some systems or extremely complex for other systems. In general the PES for chemical reactions will be complex and involve multiple electronic states and the couplings between these states. When performing dynamical simulations, one must have the ability to generate the potential and the gradient of the potential for the molecular system of interest at each nuclear configuration encountered during the course of the simulation. There exist an abundance of software packages for the evaluation of these *ab initio* stationary-state electronic energies and gradients. These software packages typically have many levels of theory that may be chosen to calculate these energies and gradients and the level chosen will ultimately depend on the computational cost that one is willing to absorb. Computational scaling factors ranging from linear scaling, $O(N)$, for the lowest level of theory to $O(N^7)$ for higher levels of theory must be spent for *each* energy evaluation, where N is the number of electrons in the system. This so-called method of direct dynamics, i.e., performing an *ab initio* energy and gradient evaluation at each time step has a considerable computational barrier for each simulation being performed. Furthermore, this computational barrier is markedly exacerbated when one must perform 10^3-10^5 simulations in order to obtain meaningful statistics of outcomes.

An alternative methodology for performing molecular dynamics simulations of a chemical system is to absorb the cost of calculating the global PES *a priori* by calculating energy points on a grid and using this grid of data as the basis for an analytic representation of the PES. This approach in no way diminishes the computational scaling factors for each *ab initio* energy evaluation ($O(N)$ to $O(N7)$); instead, once enough of a grid of energies is calculated such that the multidimensional reaction pathway is manifest, then the *ab initio* energy evaluations may cease. Then, already having the *a priori* calculation of the PES, one may perform the dynamical calculations upon this grid of data using analytic representations of the data that have scaling factors that vary according to the number of grid points, K. The following properties of the analytic representations should be expected:[2]

1.) The method ought reproduce the *ab initio* PES data from which it is formulate, i.e., the method should be an interpolant.

2.) The method ought to be completely general and easily portable to the study of very different chemical systems and, thus, very different PESs.

3.) The method ought to be computationally inexpensive. The computational cost ought to reside in the *a priori* calculation of the PES of which the method is independent. If this criteria is satisfied, then this allows for fast dynamical calculations and the possibility of meaningful statistics of outcomes.

4.) The method should not assume any regularity to the grid of *ab initio* data, i.e., features of the PES that are thought to be of extra importance to the reaction should be allowed to have higher densities of *ab initio* data.

5.) The method should require only the *a priori* calculation of the PES and not higher order derivatives. This can be particularly important since it is often the case for higher levels of *ab initio* theory that the programming of higher order derivatives are more slowly developed and their computational costs often exceed that of the evaluation of the energy.

6.) The method should generate an analytic representation of the PES that is smooth and continuous. The conservation of energy and momentum for molecular dynamics simulations will be very sensitive to this requirement.

Methodology

The multiquadric interpolant has been used as an accurate and fast global representation of multidimensional *ab initio* PESs.[2-7] This interpolant may be written as:

$$E(P) = \sum_{i=1}^{K} c_i \sqrt{d_{i,P}^2 + \Delta} \ . \qquad (1.)$$

where $d_{i,P}^2$ is the distance from point i to the point P, (point P being where the energy is being evaluated) and Δ is an *ad hoc* parameter. The interpolant analytically reproduces the underlying grid of energies used in it's formulation. As can be seen from Eq. (1) going to higher dimensionality does not significantly alter the difficulty of employing the interpolant and the interpolant will work for both uniform grids and grids of scattered data. The primary computational cost in the formulation of the interpolant involves solving a K × K system of linear equations. The cost of this evaluation is only required once for a given PES. Once one has solved for the coefficients, c_i, the interpolant may be used to evaluate the energy and higher order derivatives anywhere inside the underlying grid. The computational cost of a single energy or gradient evaluation is $O(K)$ with a very small prefactor. The primary difficulty with the utilization of this interpolant involves the *a priori* solution for the *ad hoc* parameter, Δ. As shown below, the interpolation of the energy is very stable with changes in Δ, but the accuracy of higher order derivatives may be significantly improved with optimization of this parameter. Methods have been developed[2,3] for the *a priori* solution of Δ, but they all involve solving the K × K system of linear equations multiple times.

Results

The multiquadric interpolant has recently been applied to the study of the spin-aligned Bose-Einstein condensate Li_3 system,[2,7] where an accurate and smooth representation of the *ab initio* PES is of critical importance. The global *ab initio* PES consisted of 1122 energy points calculated at the full configuration interaction (FCI) level of theory for the valence electrons and the grid of data was highly non-uniform. This is a very sparse global PES for a system of three degrees of freedom. The accuracy of the multiquadric interpolant was examined by formulating the interpolant using 1072 of the 1122 *ab initio* energies. The remaining 50 data points were used for the purposes of performing numerical tests on the accuracy of the interpolant. Table 1 shows the average, maximum, and minimum errors of the interpolated energies at these randomly

selected 50 points. The input coordinates of the PES had accuracy in the hundred-thousandth place (E-5), while the *ab initio* energies contained accuracy in the millionth place (E-6).

Table 1. Table of the average, maximum, and minimum errors in the interpolated energies at the 50 randomly scattered data points of the Li_3 PES.

Test	Avg. error (a.u.)	Max. error (a.u.)	Min. error (a.u.)
50 points	1.1E-5	1.52E-4	<E-6

As can be seen in Table 1 the interpolant has an average error of 1.1×10^{-5} Hartrees. This average error was found to be insensitive to changes in Δ. The maximum errors were found to reside at data points that were near the edge of the underlying grid of *ab initio* data where there were large gradients due to the close interactions of a Li_2 diatomic.

In addition to examining the accuracy of the energies of the multiquadric interpolant for the Li_3 PES, the accuracy of the interpolated gradients of the energy was also examined. Accuracy and smoothness of these gradients is critical for any molecular dynamics simulations since a classical Hamiltonian is often used for the propagation of the nuclei. Since these gradients were not calculated at the *ab initio* level of theory, the *ab initio* PES was fit to an analytic potential[7] and this analytic potential was used to generate a grid of data at the 1072 points where the *ab initio* energies are known. A new multiquadric interpolant was formulated using these 1072 energy points generated from the analytic potential and the gradients of this newly formulated interpolant were compared against the analytic gradients of the analytic potential. Displayed in Figure 1 is the root mean square error of the interpolated gradients along the symmetry coordinates S_1, S_2, and S_3 as a function of Δ for these 50 data points.

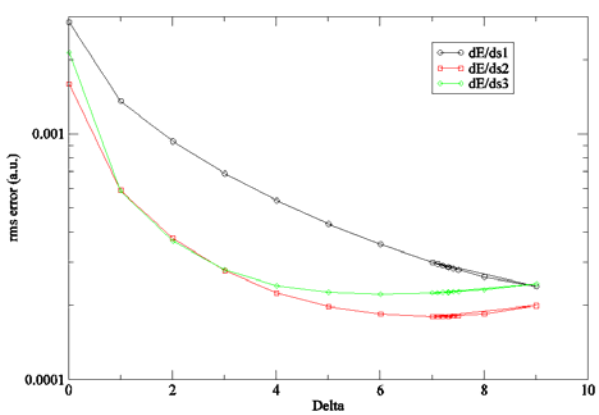

Figure 1. The root mean square error of the interpolated gradients of the energy with respect to the symmetry coordinates S_1, S_2, and S_3 as a function of the ad hoc parameter Δ.

Examining Figure 1 shows that the errors in the gradients are improved by an order of magnitude by the optimization of Δ. Figure 1 also shows that the optimized value of Δ for each of the coordinates resides at very similar values. At the optimized value of Δ, the root mean square error in the gradients at the 50 randomly selected data points ~2.0×10^{-4} Hartree/bohr along each coordinate.

Conclusions and Future

The multiquadric interpolant has been shown to be an accurate and robust method for obtaining an analytic representation of *ab initio* PESs and has further been shown to meet all of the desired properties outlined in the Introduction. The primary difficulty in the employment of the interpolant involves the *a priori* solution of the *ad hoc* parameter Δ for a given PES. Although methods have been developed for this purpose, this remains an area of active research.

Work is now being performed to utilize the multiquadric interpolant for the purposes of being able to model much larger reactive chemical systems. The goal of this work is to accurately model molecular dynamics simulations of large systems employing much higher levels of *ab initio* theory than methods of direct dynamics. As noted earlier, methods of direct dynamics, while having the benefit of using the *ab initio* PES, nevertheless often have the liability of very poor statistical outcomes. As a result, it is tempting to use levels of *ab initio* theory that may be inappropriate for systems that involve multiple electronic surfaces, which results in surface features (barriers, potential wells, etc.) that are highly over- or underestimated. Instead of applying direct dynamics, work is being performed that utilizes the *O(K)* scaling of the multiquadric interpolant to time dependent groups of the total potential for large systems.

Acknowledgement. I would like to thank Walter J. Stevens and James P. Burke for making their Li_3 PES available. I would also like to thank Union University for research support. Acknowledgement is made to the Donors of the American Chemical Society Petroleum Research Fund for partial support of this research. This research was also supported by an award from Research Corporation.

References

(1) Schatz, G. In *Reaction and Molecular Dynamics*; Lecture Notes in Chemistry; Lagana, A. and Riganelli A., Eds.; Springer, Berlin, 2000; pp. 15-32.
(2) Salazar, M. R. *Chem. Phys. Lett*. **2002**, *359*, 460.
(3) Salazar, M. R. and Bell R. *J. Comp. Chem*. **1998**, *19*, 1431.
(4) Salazar, M. R. and Simons J. *J. Chem. Phys*. **1996**, *105*, 10919.
(5) Salazar, M. R. and Simons J. *J. Chem. Phys*. **1999**, *110*, 229.
(6) Bell R. and Simons J. *J. Phys. Chem. A*. **1999**, *103*, 539.
(7) Colavecchia, F.D.; Burke, J.P. Jr; Stevens, W.J.; Salazar, M.R.; Parker, G. A.; Pack, R.T *J. Chem. Phys*. **2003**, *118*, 5484.