JCTC Journal of Chemical Theory and Computation

Non-Hermitian Multiconfiguration Molecular Mechanics

Oksana Tishchenko* and Donald G. Truhlar*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received February 13, 2009

Abstract: We present a new version of the multiconfiguration molecular mechanics (MCMM) algorithm for fitting potential energy surfaces of complex reactive systems. The main improvement consists in allowing the valence bond configuration interaction matrix to be non-Hermitian, which broadens the range of geometries over which the potential energy surface can be fit accurately. A second improvement is that the new algorithm has simpler gradients and Hessians and executes faster. The performance of the new algorithm is evaluated using the example of two model reactions.

1. Introduction

The representation of potential energy surfaces for chemical reaction dynamics continues to present a multifaceted challenge. For small systems, there has been great progress with new approaches for fitting surfaces.[1−7] For larger systems, the only practical approach is often direct dynamics; in this approach, instead of using a prefitted potential energy surface, "all required energies and forces for each geometry that is required for evaluating dynamical properties are obtained directly from electronic structure calculations."[8] This raises the cost unless inexpensive electronic structure methods (such as ab initio Hartree−Fock theory,[9,10] neglect-of-differential-overlap molecular orbital theory,[8,11] or diatomics-in-molecules valence bond theory[12]) are used. However, more reliable results can be obtained if direct dynamics calculations are based on density functional theory,[13,14] multiconfiguration ab initio wave function theory,[15] or multicoefficient correlation methods.[16] Therefore, a variety of approaches intermediate between straight direct dynamics and straight fitting have arisen, such as use of specific reaction parameters[17,18] and low-dimensionality interpolatory methods.[19,20] In the same spirit is the use of combined quantum mechanical and molecular mechanical methods[21] or methods that combine semiempirical valence bond theory[22] with molecular mechanics valence bond diabatic states, as is done in modeling ionic−covalent interactions,[23] in combining valence bond theory for reactive atoms with molecular mechanics for spectator atoms,[24] in the empirical valence bond method of Warshel and Weiss,[25] or in the approximate valence bond method of Bala et al.[26] One very promising method of the latter type is multiconfiguration molecular mechanics[27−34] (MCMM), which also combines semiempirical valence bond theory with molecular mechanics valence bond diabatic states, but in a way that in principle allows systematic improvement of potential energy surfaces to an arbitrarily high accuracy.

Multiconfiguration molecular mechanics was shown[27,28,30,31] to be successful for gas-phase kinetics when the dynamical calculations are based on variational transition state theory[35−37] with multidimensional tunneling[8,38−42] even with large-curvature tunneling approximations[8,39,42] that require energies at points significantly removed from the minimum energy path, and for liquid-phase kinetics based on umbrella sampling.[33] We have also shown that accurate VTST/MT reaction rate coefficients can be obtained using MCMM potentials constructed with as little as one or a few electronic structure Hessians,[31] using standard (e.g., MM3)[43] molecular mechanics force fields. In subsequent work,[34] by testing an MCMM potential for quasiclassical trajectories, we found that, even with relatively good molecular mechanics force fields, it is hard to converge a global potential energy surface to better than 1−2 kcal/mol. In that study, we identified the key limitation of achieving high global accuracy in practical calculations. In particular, we found that the key limitation is the inability of the previous formulation to improve the potential energy surface in regions where the accurate result is higher in energy than the lower of the reactant and product

* Corresponding author e-mail: o_t@t1.chem.umn.edu; truhlar@umn.edu.

10.1021/ct900077g CCC: $40.75
Published on Web 05/21/2009

molecular mechanics approximations. Here, we show that this can be overcome by a non-Hermitian formulation of the theory. The goal of the present work is to show that the new non-Hermitian formulation of the MCMM procedure can be used to fit semiglobal potential energy surfaces with much higher accuracy than the original Hermitian MCMM.

2. Non-Hermitian MCMM

2.1. Key Elements of the New Formulation. In MCMM, the potential energy V at a molecular geometry $\mathbf{x}$ is approximated by the lowest eigenvalue of a valence bond configuration interaction Hamiltonian matrix $\mathbf{H}$, defined by the following:

$$\mathbf{H} = \begin{pmatrix} H_{11}(\mathbf{x}) & \beta(\mathbf{x}) \\ \beta(\mathbf{x}) & H_{22}(\mathbf{x}) \end{pmatrix} \tag{1}$$

where H_{11} and H_{22} are analytical representations of valence bond configurations of the reactant and the product (e.g., molecular mechanics potentials), and β is the approximation to the off-diagonal matrix element, H_{12}. In the original MCMM, β is given by Shepard interpolation of modified Taylor series for H_{12},[27] but in non-Hermitian MCMM, we obtain the square β_o^2 of a zeroth approximation to H_{12}^2 by Shepard interpolation of the unmodified H_{12}^2, and then we write β in terms of β_o. The Shepard interpolation step yields

$$\beta_o^2(\mathbf{x}) = \sum_{k=1}^{N} w_k(\mathbf{x}) T_{12}^2(\mathbf{x}, k) \tag{2}$$

where w_k is a Shepard-interpolation weight function, and each quantity $T_{12}^2(\mathbf{x}, k)$ is a second-order Taylor series of H_{12}^2 at a geometry $\mathbf{x}_k$. In non-Hermitian MCMM, we then approximate β by the following:

$$\beta(\mathbf{x}) = \begin{cases} |\beta_o(\mathbf{x})|; & \beta_o^2(\mathbf{x}) \geq 0 \\ iu|\beta_o(\mathbf{x})|; & \beta_o^2(\mathbf{x}) < 0 \end{cases} \tag{3}$$

where

$$u(\mathbf{x}) = \begin{cases} 1; & \beta_o^2(\mathbf{x}) \geq -\Delta^2/4 \\ \Delta/(2|\beta_o|); & \beta_o^2(\mathbf{x}) < -\Delta^2/4 \end{cases} \tag{4}$$

and

$$\Delta = H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}) \tag{5}$$

There are two key points to emphasize in the above formulation. (i) First, by allowing $\mathbf{H}$ to be non-Hermitian, we broaden the range of geometries for which the MCMM fit can be accurate and thus obtain more accurate representations of potential energy surfaces. The meaning of "broadening" is the inclusion, in addition to the geometries where the true potential is lower than either H_{11} and H_{22}, of all those geometries at which the true potential is above H_{11} or H_{22} (this situation is not typical near a saddle point, but it may be the case when one interpolates a global potential energy surface). At all of these points, the MCMM potential and its first and second derivative were previously set equal to the potential and derivatives of H_{11} or H_{22}, whichever is lower, but the new MCMM formalism allows the improvement of the fitted potential energy surface at such geometries due to H_{12}. (ii) Second, by applying the cutoff function u after the interpolation rather than applying a cutoff function at each Shepard point k, as was done previously, we greatly simplify the algebra, which results in shorter computation times.

The condition $\beta_o^2 < 0$ in eq 3 corresponds to the target potential energy surface being greater than one of the diagonal elements, i.e., than one of H_{11} or H_{22}; the second row of eq 3 allows us to improve the molecular mechanics approximation in this case, and, in fact, we can make MCMM agree exactly with the target data if $\beta_o^2 \geq -\Delta^2/4$.

2.2. Details of Algorithm. The second order Taylor series expansions $T_{12}^2(\mathbf{x}, k)$ used in the Shepard interpolation of eq 2, are constructed in the same way as in steps eqs A5−A11 of the Appendix of ref 34, which is the same as in an earlier version[27] of the MCMM procedure. In particular, we define a matrix $\mathbf{H}^{(k)}$ at each Shepard node k as follows:

$$\mathbf{H}^{(k)} = \begin{pmatrix} H_{11}^{(k)} & H_{12}^{(k)} \\ H_{12}^{(k)} & H_{22}^{(k)} \end{pmatrix} \tag{6}$$

Expanding both the diagonal and the off-diagonal elements in Taylor series around a geometry $\mathbf{x}_k$ and using a Taylor series reversion[44] of H_{12}^2, one obtains,

$$T_{12}^2(\mathbf{r}, k) = D^{(k)}\left(1 + \mathbf{b}^{(k)T}\Delta\mathbf{r}^{(k)} + \frac{1}{2}\Delta\mathbf{r}^{(k)T}\mathbf{C}^{(k)}\Delta\mathbf{r}^{(k)}\right) \tag{7}$$

where $T_{12}^2(\mathbf{r}, k)$ is the value at a geometry $\mathbf{r}$ of the expansion of H_{12}^2 in a quadratic Taylor series centered at a Shepard node k; D, b, and C are Taylor series coefficients at that Shepard node k defined by eqs 20−22 of ref 27, and $\Delta\mathbf{r}^{(k)}$ is the difference between the value of a coordinate $\mathbf{r}$ at a given geometry and at Shepard node k. Note that this step and the Shepard interpolation given in eq 2 are performed in internal rather than in Cartesian coordinates $\mathbf{x}$; this set of coordinates (which is called set $\mathbf{r}$ to be consistent with the notation of a previous paper)[32] can be redundant or nonredundant.

The lowest eigenvalue of eq 1, which is an MCMM approximation to the Born−Oppenheimer potential energy, is given by

$$V = \frac{1}{2}(H_{11}(\mathbf{q}(\mathbf{x})) + H_{22}(\mathbf{q}(\mathbf{x})) - [(H_{11}(\mathbf{q}(\mathbf{x})) - H_{22}(\mathbf{q}(\mathbf{x})))^2 + 4\beta^2(\mathbf{r}(\mathbf{x}))]^{1/2}) \tag{8}$$

where β^2 is given by the square of eq 3. Note that β^2 is equal to the quantity β_o^2 obtained directly by Shepard interpolation for all positive β_o^2 and for those negative values β_o^2 that are larger than $-\Delta^2/4$. The analytical first and second derivatives of eq 8 are given in Appendix A. These derivatives involve derivatives of H_{11}, H_{22}, and β^2. The first and second derivatives of H_{11} and H_{22} are calculated analytically by a molecular mechanics code in "natural" internal coordinates $\mathbf{q}$ that are used to express molecular mechanics potentials and are then transformed to Cartesian coordinates. The first and second derivatives of β^2 are calculated from eqs 2−4 in internal coordinates $\mathbf{r}$ and are also transformed to Cartesian coordinates. The gradient and Hessian of V given by eqs 10 and 11 of the Appendix are calculated in Cartesian coordinates. The usage of different sets of coordinates in the MCMM procedure is discussed in detail in ref 32.

The weight function w_k of eq 2 is given by,

$$w_k = \frac{\dfrac{Y_k}{d_k(\mathbf{s})^4}}{\sum_{k=1}^{N+2} \dfrac{Y_k}{d_k(\mathbf{s})^4}} \quad (9)$$

where the variable d_k is a generalized distance[27] between the current geometry and the geometry at Shepard node k, expressed in internal coordinates **s** (we give the internal coordinates used in eq 9 a different name because it is usually convenient to use different coordinates here than were used above). Y_k can be approximated in different ways for different applications. For variational transition state theory calculations (as in the application presented below), where one is only interested in the region of a potential energy surface ranging, in terms of the intermolecular distance between the fragments, from the van der Waals complex of the reactants to the van der Waals complex of the products, one can simply take Y_k as unity. When one requires the potential energy surface at larger intermolecular distances, i.e., beyond the van der Waals complexes of reactants and products, then one can use a function like eq 31 of ref 34.

3. Application to Model Reactions OH + H_2 → H_2O + H and H_2O + H → OH + H_2

The improved MCMM scheme can be applied for constructing semiglobal[45] potential energy surfaces that are invariant under permutations of selected identical nuclei, where such invariance is required (for example, in full-dynamics calculations of gas-phase reactions).[34] This is achieved by properly symmetrizing the diagonal and off-diagonal elements of matrix **H**.[32] It can also be used in VTST/MT calculations that only rely on the knowledge of a potential energy surface in the vicinity of a single reaction swath in which case the nuclear permutation symmetry need not be imposed. When we perform VTST/MT calculations without enforcing the nuclear permutation symmetry, the total number of Shepard points is reduced (e.g., from $m!N$ to N for a reactive system with m low-energy symmetrically equivalent reaction channels), and this reduction results in shorter computation times.

The present application is restricted to the nonsymmetrized potential energy surface of the reactions OH + H_2 → H_2O + H and H_2O + H → OH + H_2, which are used to evaluate the performance of the improved interpolation scheme. We compare the interpolated energies and the VTST/MT rate constants calculated using the MCMM potential energy surfaces to target results calculated directly, i.e., without interpolation. Since our goal is to test an MCMM interpolation method, the use of any electronic structure method that yields a smooth potential energy surface is appropriate. As in the previous work,[34] the target results are obtained using the MPWB1K[46] density functional in conjunction with the 6-31+G(d,p)[47] basis set.

The VTST/MT calculations were performed using the POLYRATE[48] code, the MPWB1K/6-31+G(d,p) energies, gradients, and Hessians were obtained using the *Gaussian*[49] code, the MCMM energies, gradients, and Hessians were obtained using a modified version the MC-TINKER2008-2 code[50] (MC-TINKER2008-2 uses TINKER[51] for molecular mechanics calculations). The VTST/MT calculations on the interpolated surfaces were carried out using the MC-TINKERATE[52] program, which is an interface between MC-TINKER-2008-2 and POLYRATE.

Molecular mechanics force fields used in the present application are given in the Supporting Information. All parameters, except bond dissociation energies, are taken from the previous work.[34] In previous work, all bond dissociation energies were set to values exceeding their accurate values to avoid negative V_{12}^2, but now we use values that are close to experimental values, and this improves the quality of the fit even in regions where one does not place Shepard points.

Two interpolated potential energy surfaces (PES) are considered in the present paper. The first, called PES1, is based on 14 Shepard points, and the second, called PES2 is based on 11 Shepard points. In each case, there are Shepard points at the reactant and product van der Waals minima, at which locations we set $T_{12}^2 = 0$; whereas at the other Shepard points, called electronic structure Shepard points, one obtains T_{12}^2 from MPWB1K calculations. There are 12 electronic structure Shepard points for PES1 and 9 electronic structure Shepard points for PES2.

The electronic structure Shepard points for PES1 are placed at the saddle point optimized at the target level and at 11 nonstationary points. Ten nonstationary points are placed on the minimum energy reaction path (MEP) calculated at the target level, at the following locations: 0.4, 2.0, 3.5, and 4.5 kcal/mol below the saddle point on the OH + H_2 side and 0.2, 0.9, 4.9, 10.1, 14.5, and 17.1 kcal/mol below the saddle point on the H_2O + H side; and one nonstationary point is placed on the concave side of the MEP at a point where the energy is 42 kcal/mol above the H_2O + H asymptote.

The electronic structure Shepard points for PES2 are placed at the saddle point and at 8 nonstationary points. The first nonstationary point is placed on the target level MEP at the point where the energy is 2.0 kcal/mol below the saddle point on the reactant side, and the remaining 7 points were added iteratively, each on the MEP of an MCMM surface with one less electronic structure Shepard points; this is similar to the procedure described in ref 28. These points are located 0.5, 1.1, 2.5, and 2.9 kcal/mol below the saddle point on the OH + H_2 side, and 1.0, 8.0, and 8.6 kcal/mol on the H_2O + H side. The full sets of Cartesian coordinates for all Shepard points of both surfaces are given in Supporting Information.

As explained in Section 2, we use three different internal coordinate sets: set **q** for molecular mechanics calculations, set **r** for Shepard interpolation, and set **s** to calculate the Shepard weighting function. In the present application, the set **r** consists of six nonredundant internal coordinates (three bond distances, two bond angles, and a torsion), and the internal coordinates used to calculate weight function (set **s**) consists of three interatomic distances; all of these coordinates are shown in Figure 1. Previously,[32,34] we only considered cases when set **s** is equivalent to set **r**, but in the

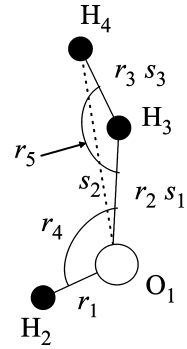

Figure 1. Internal coordinates used in Shepard interpolation (set **r**) and in eq 9 (set **s**).

present work, we introduce some flexibility in the coordinate choice by allowing **s** be different from **r**. Although it was found[27] that an optimal choice for set **s** is often not the same as for set **r** (for example, for VTST/MT calculations of an atom transfer reaction $AX\cdots B \rightarrow A\cdots XB$, a recommended choice for coordinates **s** includes three interatomic distances that undergo significant changes along the MEP, such as, $A - X$, $X - B$, and $A - B$, whereas set **r** requires at least $3N - 6$ nonredundant coordinates), the original implementation[27] of the MCMM procedure lacked the ability to correctly handle this situation. Appendix B gives the Jacobians and Hessians required for MCMM first and second derivatives for an atom transfer reaction with the local coordinates shown in Figure 1. Set **r** is also the same as the set of internal coordinates used to calculate generalized normal mode vibrational frequencies[30,37] along the minimum energy reaction path in VTST/MT calculations.

First, we consider the accuracy of the interpolated energies. In the general case $u(\mathbf{x})$ can be made smooth by joining the two regions of eq 4 by a spline function. In the present, case we simply set $u(\mathbf{x}) = 1$ because none of the geometries involved in the present tests has $\beta_o^2 < -\Delta^2/4$.

We test the surfaces on a 20×20 grid of molecular geometries. The grid was generated by varying the two key bond distances of the transferring H atom while fixing the remaining geometrical parameters (one bond distance, two angles, and a torsion) at their values at the reaction saddle point; the key distances span the ranges $r_2 = 0.78-1.92$ Å and $r_3 = 0.58-1.72$ Å, and total number of geometries is 400. Then, we deleted all energies above 64 kcal/mol with respect to $H_2O + H$; this leaves 338 points. Although these geometries do not span the whole range of dynamically important nuclear configurations, they comprise a representative set of such configurations that may be used to evaluate the accuracy of the fit. Figures 2 and 3 show two-dimensional slices of the target potential energy surface and one of the interpolated ones (PES1) as functions of these two bond distances. Table 1 lists mean unsigned errors for PES1 and PES2 calculated using an old MCMM algorithm[27,34] and using the improved MCMM procedure presented above. The errors are shown as functions of potential energy for four different energy ranges below 64 kcal/mol; the smallest subset of geometries (in the lowest energy range) comprises 46 geometries, and the largest subset comprises 338 geometries. The results indicate that the new MCMM procedure leads to more accurate interpolated potential energies than the Hermitian one, especially in the regions where one places Shepard points. The more accurate fit for PES1 as compared to PES2 is due to the placement of Shepard points in a wider energy range in the former case. In particular, the energy

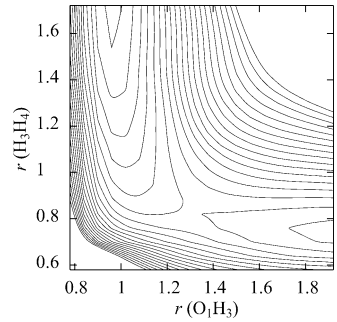

Figure 2. Equipotential contour plot of PES1 as a function of the OH and HH distances. The remaining internal coordinates (one bond distance, two bond angles, and a torsion) are fixed at their values at the reaction saddle point. Contours start at −10.0 kcal/mol and are equally spaced by 3 kcal/mol, with the zero of energy at the OH + H_2 asymptote. Bond distances are in Å.

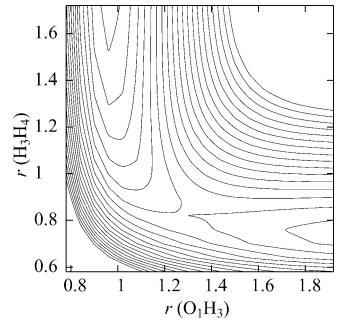

Figure 3. Equipotential contour plot of the target (uninterpolated) potential energy surface. The remaining internal coordinates (one bond distance, two bond angles, and a torsion) are fixed at their values at the reaction saddle point. Contours start at −10.0 kcal/mol and are equally spaced by 3 kcal/mol, with the zero of energy at the OH + H_2 asymptote. Bond distances are in Å.

Table 1. Mean Unsigned Errors[a] in energies for Two Potential Energy surfaces for Various Ranges of Potential energy[b]

energy range	N[c]	PES1 (current algorithm)	PES2 (current algorithm)	PES1 (old algorithm)	PES2 (old algorithm)
<14	46	0.1	0.5	0.6	1.0
<20	116	0.3	0.7	0.6	0.8
<26	177	0.4	0.9	0.7	0.9
<64	338	0.9	1.9	1.2	1.6

[a] In kca/mol. [b] Zero of energy for this table corresponds to H_2O + H. The forward and reverse barrier heights for this reaction with full MPWB1K/6-31+G(d,p) calculations are 4.94 and 18.57 kcal/mol, respectively. Therefore, the number in the first column can be converted to a scale relative to reactants by subtracting 18.57−4.94 = 13.63 kcal/mol. [c] Number of points in this range.

Table 2. Mean Unsigned Percentage Errors in Rate Coefficients for Two Potential Energy Surfaces

T	k^{CVT}	$k^{CVT/ZCT}$	$k^{CVT/SCT}$	$k^{CVT/LCT}$	$k^{CVT/\mu OMT}$
			PES1		
			$OH + H_2 \rightarrow HOH + H$		
300	8	13	3	15	3
400	7	10	5	11	5
600	6	9	7	9	7
			$HOH + H \rightarrow OH + H_2$		
300	7	13	2	14	2
400	6	9	4	10	4
600	4	8	6	8	6
			PES2		
			$OH + H_2 \rightarrow HOH + H$		
300	0	1	10	2	10
400	1	4	2	4	2
600	2	4	1	4	1
			$HOH + H \rightarrow OH + H_2$		
300	1	1	11	1	11
400	0	3	3	3	3
600	1	3	0	3	0

Table 3. Mean Unsigned Percentage Errors Averaged over Three Temperatures (300 K, 400 K, and 600 K) for Reactions $OH + H_2 \rightarrow HOH + H$ and $HOH + H \rightarrow OH + H_2$

T	k^{CVT}	$k^{CVT/ZCT}$	$k^{CVT/SCT}$	$k^{CVT/LCT}$	$k^{CVT/\mu OMT}$
PES1	6	10	5	11	5
PES2	1	3	5	3	5

ranges where the Shepard points are present are 1.5−42.0 kcal/mol for PES1 and 10.0−18.6 kcal/mol for PES2.

Tables 2 and 3 present mean unsigned percentage errors in rate coefficients calculated using the two new MCMM potential energy surfaces. The mean unsigned percentage error is defined as percentage accuracy of the rate coefficient calculated using an MCMM potential as compared to the target result[28,31] obtained via direct dynamics calculations. While the canonical variational transition state rate constant, k^{CVT}, provides a test of the accuracy of the fit near the top of the free energy of activation barrier, the rate coefficients including the tunneling correction (ZCT,[53] SCT,[40,53] LCT,[8] and μOMT)[8,42] depend on the potential in a much wider region. The results for k^{CVT} and for other rate constants are on the average better for PES2 than are the results for PES1. The smaller errors for PES2 are mainly due to the nearly precise location of the variational transition state on this surface because of the availability of a Shepard point in a close proximity to the reaction bottleneck (note that unlike PES1, the PES2 is constructed by placing Shepard points according to a scheme similar to the one in ref 28 particularly designed for VTST/MT calculations, whereas the locations of the data points on PES1 are more or less arbitrary). Excellent results on both interpolated surfaces are obtained for k^{LCT} calculations, which are very sensitive to the potential off the MEP, even though in the case of PES2 no Shepard points are placed on the concave side of the minimum energy reaction path and in the case of PES1 only one point is placed there.

The MCMM scheme is a method for constructing full-dimensional potential energy surfaces for dynamics applications (such as VTST, classical trajectories, or other methods). Although the full advantages of the improved MCMM technique are most apparent in full dynamics simulations, the results of VTST/MT reaction rate calculations also benefit from the improvements introduced the present work in the following ways: (i) the lowest diagonal matrix element may appear to be above the true potential energy surface even at geometries on or near MEP, therefore, using a non-Hermitian matrix **H** improves the accuracy of the MEP on the fitted surface, (ii) the VTST/MT calculations, especially those including large-curvature tunneling, depend on the surface in a much wider range of geometries, therefore, by allowing **H** to be non-Hermitian one improves the overall results, and (iii) eliminating artificial modification in Shepard interpolation introduced in the original MCMM[27] also eliminates sudden changes in the coupling term (caused by forcing $T^2_{\uparrow 2}$ to zero at geometries where is negative) and thus results in a smoother MEP and smoother vibrationally adiabatic ground-state potential curve (obtained by adding zero-point energies of orthogonal harmonic vibrations to the MEP), as well as smoother changes in vibrational frequencies along the reaction path.

Note that H_{11} and H_{22} are expressed here in terms of standard molecular mechanics functions, but one could also replace them with very accurate fits to ab initio data if higher precision in the calculations is sought. Even when H_{11} and H_{22} are represented by standard molecular mechanics, the results in Table 3 indicate that the rate coefficients are converged within 5% of the target results (when one uses the strategy for placing Shepard points designed for VTST calculations,)[28] or within 15% (when the points are placed more or less arbitrarily); in each case these results are within the typical uncertainty of 25% in best estimates of rate coefficients. This implies that MCMM can be used in an automated way (without fitting or adjustments and using a prescription[28] for placing Shepard points) to construct reasonably accurate semi-global representations of potential energy surfaces. However, due to the possibility to replace the standard molecular mechanics picture with more accurate representations for H_{11} and H_{22}, there is enough flexibility to get an arbitrarily high accuracy.

Although the new algorithm is tested here for a tetratomic system, because that allows for well-defined high-precision tests of its ability to fit the dependence on reactive coordinates, we emphasize that MCMM is designed not to compete with modern algorithms for fitting few-body surfaces to spectroscopic accuracy (although, when used with care, it may be competitive with such algorithms) but rather is designed to provide a practical method to fit potential energy surfaces for complex systems.

4. Summary

In this work, we present a new formulation of the multi-configuration molecular mechanics algorithm that improves the accuracy of interpolated potential energy surfaces.

Acknowledgment. This work was supported in part by the NSF under Grant No. CHE07-04974 (dynamics of complex systems, global potential energy surfaces), by the DOE (gas-phase variational transition state theory) under Grant No. CHE07-04974, and by the office of Naval Research (integrated software tools for dynamics) under award No. N00014-05-1-0538.

Supporting Information Available: Force field parameters and geometries used to construct potential energy surfaces and reaction rate coefficients for reactions OH + $H_2 \rightarrow H_2O$ + H and H_2O + H → OH + H_2. This material is available free of charge via the Internet at http://pubs.acs.org.

Appendix A: The Gradient and Hessian of *V*(x) in Cartesian Coordinates

The components of the gradient of an MCMM approximation of the potential energy are given by,

$$G_i = \frac{\partial V}{\partial x_i} = \frac{1}{2}\left(G_{11i}(\mathbf{x}) + G_{22i}(\mathbf{x}) - \left(\frac{2\left(\frac{\partial \beta^2(\mathbf{r}(\mathbf{x}))}{\partial x_i}\right) + (H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))(G_{11i}(\mathbf{x}) - G_{22i}(\mathbf{x}))}{((H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))^2 + 4\beta^2(\mathbf{r}(\mathbf{x})))^{1/2}}\right)\right) \tag{10}$$

where H_{11}, H_{22} are diagonal elements of matrix **H**, G_{11i}, and G_{22i} are components of their gradients, and β is given by eq 3. Note that the first and second derivatives of H_{11} and H_{22} are calculated in internal coordinates **q**(**x**) and then transformed to Cartesian coordinates. The elements of the Hessian of an MCMM approximation of the potential energy are given by,

$$F_{ij} = \frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{1}{2}(F_{11ij}(\mathbf{x}) + F_{22ij}(\mathbf{x})) + \frac{\left(2\left(\frac{\partial \beta^2(\mathbf{r}(\mathbf{x}))}{\partial x_i}\right) + (H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))(G_{11i}(\mathbf{x}) - G_{22i}(\mathbf{x}))\right)}{((H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))^2 + 4\beta^2(\mathbf{r}(\mathbf{x})))^{3/2}} \times \frac{\left(2\left(\frac{\partial \beta^2(\mathbf{r}(\mathbf{x}))}{\partial x_j}\right) + (H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))(G_{11j}(\mathbf{x}) - G_{22j}(\mathbf{x}))\right)}{((H_{11}(\mathbf{x}) - H_{22}(\mathbf{x}))^2 + 4\beta^2(\mathbf{r}(\mathbf{x})))^{3/2}} - \frac{(G_{11i} - G_{22i})(G_{11j} - G_{2j})}{((H_{11} - H_{22})^2 + 4\beta^2(\mathbf{r}(\mathbf{x})))^{1/2}} - \frac{2\left(\frac{\partial^2 \beta^2(\mathbf{r}(\mathbf{x}))}{\partial x_i \partial x_j}\right) + (H_{11} - H_{22})(F_{11ij} - F_{22ij})}{((H_{11} - H_{22})^2 + 4\beta^2(\mathbf{r}(\mathbf{x})))^{1/2}} \tag{11}$$

where F_{11ij} and F_{22ij} are the elements of Hessians of H_{11} and H_{22}. The first and second derivatives of β^2 with respect to the coordinates **r** are the same as the derivatives of β_o for all $\beta_o^2 \geq -\Delta^2/4$. These derivatives are given by

$$\mathbf{g} \equiv \frac{\partial \beta_o^2}{\partial \mathbf{r}} = \sum_{k=1}^{N}\left[\frac{\partial w_k}{\partial \mathbf{r}} T_{12}^2(\mathbf{r};k) + w_k D^{(k)}(\mathbf{b}^{(k)} + \mathbf{C}^{(k)}\Delta \mathbf{r}^{(k)})\right] \tag{12}$$

$$\mathbf{f} \equiv \frac{\partial^2 \beta_o^2}{\partial \mathbf{r}^2} = \sum_{k=1}^{N}\left(\frac{\partial^2 w_k}{\partial \mathbf{r}^2} T_{12}^2(\mathbf{r},k) + \frac{\partial w_k}{\partial \mathbf{r}}\mathbf{g}(\mathbf{r}) + D^{(k)}(\mathbf{b}^{(k)} + \mathbf{C}^{(k)}\Delta r^{(k)})\left(\frac{\partial w_k}{\partial \mathbf{r}}\right)^T + w_k D^{(k)}\mathbf{C}^{(k)}\right) \tag{13}$$

where

$$\frac{\partial w_k}{\partial r_\alpha} = \sum_{\gamma=1}^{\Gamma} \frac{\partial w_k}{\partial s_\gamma}\frac{\partial s_\gamma}{\partial r_\alpha} \tag{14}$$

$$\frac{\partial^2 w_k}{\partial r_\alpha \partial r_\beta} = \sum_{\gamma=1}^{\Gamma}\sum_{\gamma=1'}^{\Gamma} \frac{\partial s_\gamma}{\partial r_\alpha}\frac{\partial^2 w_k}{\partial s_\gamma \partial s_{\gamma'}}\frac{\partial s_{\gamma'}}{\partial r_\beta} + \sum_{\gamma=1}^{\Gamma}\frac{\partial w_k}{\partial s_\gamma}\frac{\partial^2 s_\gamma}{\partial r_\alpha \partial r_\beta} \tag{15}$$

where **r** and **s** are the sets of the internal coordinates used in Shepard interpolation and in calculations of the weight function.

Appendix B: Internal Coordinates s and r Used in the Present Application and the Jacobians and Hessians Required by eqs 14 and 15

The internal coordinates **s** and **r** used in eqs 7 and 9 are $\mathbf{r} \equiv \{r_1, r_2, r_3, r_4, r_5, r_6\}$ and $\mathbf{s} \equiv \{s_1, s_2, s_3\}$, respectively; these coordinates are shown in Figure 1. The Jacobians required by eqs 14 and 15 for these coordinates are

$$\frac{\partial s_1}{\partial \mathbf{r}} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \frac{\partial s_2}{\partial \mathbf{r}} = \begin{bmatrix} 0 \\ \dfrac{r_2 - r_3\cos(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} \\ \dfrac{r_3 - r_2\cos(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} \\ 0 \\ \dfrac{r_2r_3\sin(r_5)}{r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} \\ 0 \end{bmatrix} \quad \frac{\partial s_3}{\partial \mathbf{r}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{16}$$

The corresponding matrices of second derivatives are as follows:

$$\frac{\partial^2 s_1}{\partial \mathbf{r}^2} = \mathbf{0} \tag{17}$$

$$\frac{\partial^2 s_2}{\partial \mathbf{r}^2} = \begin{bmatrix} 0 & & & & & \\ 0 & \dfrac{1}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & & & & \\ & -\dfrac{1}{2}\dfrac{(r_2 - r_3\cos(r_5))^2}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & & & & \\ 0 & \dfrac{-\cos(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & \dfrac{1}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & & & \\ & -\dfrac{(2r_2 - 2r_3\cos(r_5))(2r_3 - 2r_2\cos(r_5))}{4(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & -\dfrac{1}{2}\dfrac{(r_3 - r_2\cos(r_5))^2}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & & & \\ 0 & 0 & 0 & 0 & & \\ 0 & \dfrac{r_3\sin(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & \dfrac{r_2\sin(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & 0 & \dfrac{r_2r_3\cos(r_5)}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{1/2}} & \\ & -\dfrac{r_2r_3\cos(r_5)(r_2 - r_3\sin(r_5))}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & -\dfrac{r_2r_3\cos(r_5)(r_3 - r_2\sin(r_5))}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & & -\dfrac{(r_2 - r_3\sin(r_5))}{(r_2^2 + r_3^2 + 2r_2r_3\cos(r_5))^{3/2}} & \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{18}$$

$$\frac{\partial^2 s_3}{\partial \mathbf{r}^2} = \mathbf{0} \tag{19}$$

where **0** is a null matrix. Note that eq (18) shows only the lower triangular part of the symmetric Hessian matrix.

References

(1) Ishtwan, J.; Collins, M. A. *J. Chem. Phys.* **1994**, *100*, 8080.

(2) Collins, M. A. *Theor. Chem. Acc.* **2002**, *108*, 313.

(3) Jin, Z.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* **2006**, *220*, 1569.

(4) Xie, D.; Xu, C.; Ho, T.-S.; Rabitz, H.; Lendvay, G.; Lin, S. Y.; Guo, H. *J. Chem. Phys.* **2007**, *126*, 074315.

(5) Dawes, R.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. *J. Chem. Phys.* **2008**, *128*, 084107.

(6) Albu, T. V.; Espinosa-Garcia, J.; Truhlar, D. G. *Chem. Rev.* **2007**, *107*, 5101.

(7) Valiron, P.; Wernli, M.; Favre, A.; Wiesenfeld, L.; Rist, C.; Kedzuch, S.; Noga, J. *J. Chem. Phys.* **2008**, *129*, 134306.

(8) Liu, L.-P.; Lu, D.-h.; González-Lafont, A.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 7806.

(9) Leforestier, C. *J. Chem. Phys.* **1978**, *68*, 4406.

(10) Baldridge, K.; Gordon, M. S.; Steckler, R.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 5107.

(11) Wang, I. S. Y.; Karplus, M. *J. Am. Chem. Soc.* **1973**, *95*, 8160.

(12) Truhlar, D. G.; Duff, J. W.; Blais, N. C.; Tully, J. C.; Garrett, B. C. *J. Chem. Phys.* **1982**, *77*, 764.

(13) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(14) Espinosa-Garcia, J.; Corchado, J. C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1997**, *119*, 9841.

(15) Garrett, B. C.; Koszykowski, M. L.; Melius, C. F.; Page, M. *J. Phys. Chem.* **1990**, *94*, 7096.

(16) Ellingson, B. A.; Pu, J.; Lin, H.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2007**, *111*, 11706.

(17) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618.

(18) Pu, J.; Truhlar, D. G. *J. Chem. Phys.* **2002**, *116*, 1468.

(19) Corchado, J. L.; Coitiño, E. L.; Chuang, Y.-Y.; Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 2424.

(20) Ruiz-Pernia, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B* **2006**, *110*, 17663.

(21) *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998.

(22) (a) London, F. *Z. Elektrochem.* **1929**, *35*, 551. (b) Eyring, H.; Polanyi, M. *Naturwissenschaften* **1930**, *18*, 914.

(23) (a) Coulson, C. A.; Danielsson, U. *Ark. Fys.* **1954**, *8*, 239. (b) Janev, R. K.; Radulovic, Z. M. *Phys. Rev. A* **1978**, *17*, 889.

(24) Raff, L. M. *J. Chem. Phys.* **1974**, *60*, 2222.

(25) (a) Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218. (b) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523.

(26) Bala, P.; Grochowski, P.; Nowinski, K.; Lesyng, B.; McCammon, J. A. *Biophys. J.* **2000**, *79*, 1253.

(27) Kim, Y.; Corchado, J. C.; Villa, J.; Xing, J.; Truhlar, D. G. *J. Chem. Phys.* **2000**, *112*, 2718.

(28) Albu, T. V.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 8465.

(29) Truhlar, D. G. *J. Phys. Chem. A* **2002**, *106*, 5048.

(30) Lin, H.; Pu, J.; Albu, T. V.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 4112.

(31) Tishchenko, O.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13530.

(32) Tishchenko, O.; Truhlar, D. G. *J. Chem. Theor. Comp.* **2007**, *3*, 938.

(33) Higashi, M.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 1032.

(34) Tishchenko, O.; Truhlar, D. G. *J. Chem. Phys.* **2009**, *130*, 024105.

(35) Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **1979**, *70*, 1593.

(36) Garrett, B. C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1979**, *101*, 4534.

(37) Jackels, C. F.; Gu, Z.; Truhlar, D. G. *J. Chem. Phys.* **1995**, *102*, 3188.

(38) Truhlar, D. G.; Isaacson, A. D.; Skodje, R. T.; Garrett, B. C. *J. Phys. Chem.* **1982**, *86*, 2252. . **1983**, *87*, 4554 (E).

(39) Truhlar, D. G.; Gordon, M. S. *Science* **1990**, *249*, 491.

(40) Liu, Y.-P.; Lynch, G. C.; Truong, T. N.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 2408.

(41) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *J. Phys. Chem.* **1996**, *100*, 12771.

(42) Fernández-Ramos, A.; Truhlar, D. G. *J. Chem. Phys.* **2001**, *114*, 1491.

(43) Allinger, N. C.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1998**, *111*, 8551.

(44) Chang, Y. T.; Miller, W. H. *J. Phys. Chem.* **1990**, *94*, 5884.

(45) When referring to MCMM potential energy surfaces, the term "global" indicates that such a surface is defined everywhere in the nuclear configuration space, whereas "semiglobal" is used to emphasize that it is aimed to represent an accurate fit in dynamically relevant regions. For example, in the application described in the present work, the reaction asymptote with all four atoms of the OH + H_2 system separated apart is not neccesserily well represented in the present fit, but because of the high energy of such configuration, this limitation does not affect dynamics in the energy range of interest.

(46) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.

(47) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(48) Corchado, J. C.; Chuang, Y.-Y.; Fast, P. L.; Villa, J.; Hu, W.-P.; Liu, Y.-P.; Lynch, G. C.; Nguyen, K. A.; Jackels, C. F.; Melissas, V. S.; Lynch, B. J.; Rossi, I.; Coitiño, E. L.; Fernández-Ramos, A.; Pu, J.; Albu, T. V.; Steckler, R.; Garrett, B. C.; Isaacson, A. D.; Truhlar, D. G. *Polyrate 9.7*; University of Minnesota: Minneapolis, MN, 2007.

(49) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(50) Tishchenko, O.; Higashi, M.; Albu, T. V.; Corchado, J. C.; Kim, Y.; Villà, J.; Xing, J.; Lin, H. Truhlar, D. G. *MC-TINKER-2008-2*; University of Minnesota: Minneapolis, MN, 2008.

(51) Ponder, J. W. *TINKER-Version 4.2*; Washington University: St. Louis, MO, 2004.

(52) Albu, T. V.; Corchado, J. C.; Kim, Y.; Villa, J.; Xing, J.; Lin, H.; Tishchenko, O.; Higashi, M.; Truhlar, D. G. *MC-TINKERATE-2008*; University of Minnesota: Minneapolis, MN, 2008.

(53) Lu, D.-H.; Truong, T. N.; Melissas, V. S.; Lynch, G. C.; Liu, Y.-P.; Garrett, B. C.; Steckler, R.; Isaacson, A. D.; Rai, S. N.; Hancock, G. C.; Lauderdale, J. G.; Joseph, T.; Truhlar, D. G. *Comput. Phys. Commun.* **1992**, *71*, 235.

CT900077G