Potential energy surfaces for polyatomic reactions by interpolation with reaction path weight: $CH_2OH^+ \rightarrow CHO^+ + H_2$ reaction

Young Min Rhee and Tae Geol Lee
Department of Chemistry and Center for Molecular Catalysis, Seoul National University, Seoul 151-742, Korea

Seung C. Park[a)]
Department of Chemistry, Science Campus, Sung Kyun Kwan University, Seoul 110-745, Korea

Myung Soo Kim[a)]
Department of Chemistry and Center for Molecular Catalysis, Seoul National University, Seoul 151-742, Korea

(Received 30 July 1996; accepted 4 October 1996)

An improved algorithm to construct molecular potential energy surfaces for polyatomic reactions is presented. The method uses the energies, gradients, and Hessians, which can be obtained from *ab initio* quantum chemical calculations. The surface is constructed by interpolating the local quadratic surfaces with reaction path weights. The method is tested with a five-atom reaction system for which an analytic potential energy surface has been reported together with classical trajectory results. An excellent agreement is achieved for energy partitioning in products obtained by trajectory calculation on the original analytic and interpolated surfaces. Reduction of error caused by the use of the reaction path weight is explained. © *1997 American Institute of Physics.* [S0021-9606(97)02402-1]

I. INTRODUCTION

Potential energy surfaces (PESs) form the operational starting point for essentially all dynamical calculations.[1] There have been numerous attempts to represent PESs for polyatomic reactions as accurately as possible and numbers of successes have been reported for three- and four-atom systems.[2–10] The traditional approach is to calculate the energies for a large number of geometrical configurations and to fit an analytic function to these energies. For reaction systems containing five or more atoms, such an approach becomes a formidable task because of the high dimensionality. Hence, a semiempirical surface[1] is usually constructed[11–15] utilizing experimental and/or *ab initio* information at a relatively small number of points. The effort needed in this less rigorous and less accurate approach is still formidable because quite a few parameters are needed for the analytic expression for the potential energy surfaces of many-atom systems. Fitting these parameters is very time consuming. Moreover, one can never be sure that an analytic expression obtained by adjusting these parameters would be globally adequate.

To circumvent the problem associated with the construction of an analytic PES, an *ab initio* direct dynamics method was proposed.[16,17] In computing classical trajectories for many-atom systems, the force on the nuclei, that is, the gradient of the PES, is required. Since the gradient can be calculated efficiently with quantum chemical methods,[18,19] the classical trajectory calculation can be done without explicit formulation of the PES. The main drawback of this method is the burden of the quantum chemical calculation needed. Hence, the quantum chemical calculation has been limited to a low level to economize on the computations. Helgaker, Uggerud, and Jensen[20] proposed the trust-region dynamics method to reduce the amount of *ab initio* calculation in this approach. A second-order model surface is constructed around a reference point, and it is remodeled only if the trajectory reaches the boundary of a trust radius. Although the energy, gradient, and Hessian are calculated only once in that radius, the calculation is still onerous because the size of the radius has to be rather small. In their application to the H_2 elimination reaction from CH_2OH^+, calculation of only one trajectory starting from the transition state was reported.[20,21] Chen, Hase, and Schlegel[22] introduced the predictor and corrector steps to the classical trajectory calculation algorithm. The local surface was represented by a fifth-order polynomial fitted to the energies, gradients, and Hessians at these step points. This allows considerably larger step sizes, significantly reducing the calculation cost. However, tens of *ab initio* calculations are still required for each trajectory, with the cost increasing in proportion to the number of trajectories.

The excessive computation time needed for the *ab initio* direct dynamics suggests that development of a method to construct a molecular PES directly from *ab initio* results, namely without using fitting parameters, is needed for dynamical study at the current level of computing power. As mentioned above, the main difficulty with this approach is the tremendously large number of *ab initio* calculations needed to characterize the PES of many-atom systems. In this regard, it is worthwhile to note that suggestions have

a)Authors to whom correspondence should be addressed.

been made to model the PES around the minimum energy path. For example, Miller and co-workers[23,24] proposed the reaction path Hamiltonian in which the potential is approximated by a series of local harmonic surfaces defined by the energy and its derivatives on this path. This formalism has been further developed and applied to various dynamical studies such as in the calculation of rate constants with the variational transition state theory.[25–29] The formalism, however, is still not adequate for the calculation of the classical trajectories or quantum dynamics.

Recently, Ischtwan and Collins[30] reported an algorithm to construct an analytic PES via interpolation of local surfaces approximated from *ab initio* calculations at various configuration points. Even though the construction started initially with the points on the minimum energy path, the PES was upgraded with the classical trajectory results. Hence, in addition to the fact that a continuous analytic PES is available, a reliable representation of the PES, at a point with a substantial displacement from the minimum energy path where the anharmonicity would be significant, is possible with this algorithm. The algorithm was tested for four-atom systems such as NH + H_2 (Ref. 30) and OH + H_2 (Ref. 31–33). Nguyen, Rossi, and Truhlar[34] also investigated the PES for the OH + H_2 system adopting the algorithm of Ischtwan and Collins and suggested that the method may be generally applicable to larger systems.

In this article, we present the results from our effort to apply and improve the Ischtwan–Collins (IC) algorithm for the dynamical investigation of five or more atom systems. The reaction system $CH_2OH^+ \rightarrow CHO^+ + H_2$ was taken as the model because both the pertinent analytic PES and classical trajectory results are available from our previous investigation.[15] The method to improve the weighting function by borrowing from the concept of the intrinsic reaction path[2,35–37] will be presented. Classical trajectory results obtained with various internal coordinate systems will be compared. Finally, the trajectory results from the interpolated PES will be compared with those on the original analytic PES to demonstrate that the former is entirely adequate for dynamical investigation.

II. METHOD

A. Modified Taylor expansion of the potential and internal coordinate systems

Following the IC algorithm, the potential in the vicinity of any point $\mathbf{R}^k$ is approximated by the second-order Taylor expansion

$$V^k(\mathbf{R}) = V(\mathbf{R}^k) + (\mathbf{R}-\mathbf{R}^k)^T \cdot \mathbf{g} + \tfrac{1}{2}(\mathbf{R}-\mathbf{R}^k)^T \cdot \mathbf{H} \cdot (\mathbf{R}-\mathbf{R}^k). \quad (1)$$

To guarantee the translational and rotational invariances of the potential, it should be expanded in terms of the internal coordinates. Since gradients ($\mathbf{g}$) and Hessians ($\mathbf{H}$) are efficiently calculated in Cartesian coordinates $\{\mathbf{x}\}$,[19] these quantities are originally obtained in Cartesians. Then they are transformed into quantities in internal coordinates by using the relations

$$\frac{\partial V}{\partial R_i} = \sum_k J_{ik} \frac{\partial V}{\partial x_k}, \quad (2)$$

$$\frac{\partial^2 V}{\partial R_i \partial R_j} = \sum_{k,l} J_{ik} J_{jl} \frac{\partial^2 V}{\partial x_k \partial x_l} + \sum_k \frac{\partial J_{ik}}{\partial R_j} \frac{\partial V}{\partial x_k}, \quad (3)$$

where J_{ij} denotes the Jacobian factor $\partial x_j / \partial R_i$. The transformation matrix element J_{ij} is well-defined only if a nonredundant coordinate system is used for $\{\mathbf{R}\}$. Such a coordinate redundancy arises for the five or more atom systems when all the interatomic distances are taken as the internal coordinates as adopted by Ischtwan and Collins in their study of four-atom systems. The *Z*-matrix type internal coordinate system[18] which is frequently used for inputs of *ab initio* calculations of polyatomic molecules is adopted here. This is a good choice for a many-atom system,[38] since it is nonredundant and easy to use. Its components are bond lengths, bond angles, and dihedral angles. Following the work[30–32] of Collins and co-workers, inverses of bond lengths rather than bond lengths themselves are used for $\{\mathbf{R}\}$. To ensure the periodicity with respect to the dihedral angles, the potential is expanded in terms of their sine functions. Then the expansion becomes

$$V^k(\mathbf{R}) = V(\mathbf{R}^k) + \mathbf{\Delta}^{kT} \cdot \mathbf{g}_\Delta + \frac{1}{2} \mathbf{\Delta}^{kT} \cdot \mathbf{H}_\Delta \cdot \mathbf{\Delta}^k; \quad (4)$$

$\mathbf{g}_\Delta$ and $\mathbf{H}_\Delta$ are the gradient and Hessian, respectively, with respect to $\mathbf{\Delta}$, where

$$\Delta_j^k = \begin{cases} \dfrac{1}{r_j} - \dfrac{1}{r_j^k}, & (r_j : \text{bond length}), \\ \theta_j - \theta_j^k, & (\theta_j : \text{bond angle}), \\ \sin(\phi_j - \phi_j^k), & (\phi_j : \text{dihedral angle}). \end{cases}$$

B. Interpolation with the reaction path weight

In the original development of the IC algorithm, the potential energy at any configuration was given by a modified Shepard interpolation, a weighted average of the Taylor series about all N points where the energies, gradients, and Hessians were known.

$$V(\mathbf{R}) = \sum_k^N w_k(\mathbf{R}) V^k(\mathbf{R}). \quad (5)$$

Here, $w_k(\mathbf{R})$ is an appropriate weighting function. As will be seen in Sec. IV A, the above scheme did not result in the efficient convergence of the trajectory results for the CH_2OH^+ system. A much better convergence was achieved by adding the reaction path concept to the weighting function. The algorithm modified for this purpose is described as follows.

Let us first classify the configurations with known energies, gradients, and Hessians in terms of their relative positions on the intrinsic reaction coordinate (IRC).[37] Namely, a

set $P_i=\{\mathbf{R}^{ij}\}$ $(j=1,\cdots,N_i)$ is the collection of configurations in the neighborhood of the point on IRC with the reaction coordinate value s_i. The first element of this set, $\mathbf{R}^{i1}$, is the point on IRC. Let us suppose that we have N_s such sets. The potential energy at an arbitrary configuration $\mathbf{R}$ is taken as a weighted average of Eq. (4) about all the points, as was done in the original IC algorithm. In the present algorithm, however, two types of weighting functions will be used, one (A_{ij}) is involved in the interpolation of functions within a given set and the other(B_i) is in interset contributions. Namely,

$$V(\mathbf{R})=\sum_i^{N_s} B_i \sum_j^{N_i} A_{ij} V^{ij}(\mathbf{R}). \tag{6}$$

Both A_{ij} and B_i are normalized weighting functions, A_{ij} being similar to the one considered in the original IC algorithm and B_i being the reaction path weight. In other words, the potential is given by the weighted average of the local interpolant, in which the extent of the displacement along the IRC is taken into account with the reaction path weight.

The normalized weight A_{ij} is defined as

$$A_{ij}=\frac{a_{ij}}{\Sigma_{j=1}^{N_i} a_{ij}}, \tag{7}$$

where the unnormalized factor a_{ij} has the following properties:[30]

$$\begin{aligned} a_{ij} &\rightarrow 0 \quad \text{as} \quad |\mathbf{R}-\mathbf{R}^{ij}| \rightarrow \infty, \\ a_{ij} &\rightarrow \infty \quad \text{as} \quad |\mathbf{R}-\mathbf{R}^{ij}| \rightarrow 0. \end{aligned} \tag{8}$$

Unlike the four-atom systems investigated by Ischtwan and Collins, the five-atom system considered here requires the use of some bond angles and dihedral angles as internal coordinates. In the calculation of weights, however, we propose to use the interatomic distances only, because the inclusion of angles sometimes produces rather unrealistic results. The simplest form of a_{ij} that is consistent with the above properties is the inverse power function of the distance,

$$a_{ij}=\frac{1}{|\mathbf{R}-\mathbf{R}^{ij}|^{2p}}. \tag{9}$$

The normalized reaction path weight B_i is defined similarly

$$B_i=\frac{b_i}{\Sigma_{i=1}^{N_s} b_i}, \tag{10}$$

where

$$b_i=\frac{1}{|\mathbf{R}-\mathbf{R}^{i1}|^{2p}}. \tag{11}$$

To guarantee the convergence of the interpolated PES, $2p$ should be larger than $3N-6$.[30] Using an excessively large $2p$ value, corresponding to a very sharp weighting function, causes some undesirable effects on the PES. This will be described in detail in Sec. IV C.

C. Data set construction

The initial N_s points on the reaction path are chosen from the configurations on IRC. Typically, many tens of configurations are needed such that the interpolated PES represents other points on the IRC adequately in energies, gradients, and Hessians. A minimum number of points can be selected as follows. At first, a relatively small number of configurations (for example, the reactant, the products, and the transition state) is selected, and the IRC is recalculated on the interpolated surface derived from this limited initial set. The recalculated IRC is compared with the original one and more points are added in the regions where the deviation is large. This procedure is repeated until the difference between the recalculated and the original IRCs falls within an acceptable level of error.

Data points off the IRC are added successively by modifying the trajectory simulation method proposed by Collins *et al.*[30,32] For a given data set (initially, this consists of data points only on the IRC), a number of trajectories is evaluated. Throughout these trajectories, configurations are sampled periodically, producing N_t trajectory points denoted as $\mathbf{R}(k)$. These points are then weighted using both the already existing points and the newly acquired trajectory points as follows:

$$h[\mathbf{R}(k)]=\frac{1}{N_t-1}\sum_{\substack{n\neq k \\ n=1}}^{N_t} \frac{v_k[\mathbf{R}(n)]}{\Sigma_{i=1}^{N_s}\Sigma_{j=1}^{N_i} v_{ij}[\mathbf{R}(k)]}, \tag{12}$$

where

$$v_k[\mathbf{R}(n)]=\frac{1}{|\mathbf{R}(n)-\mathbf{R}(k)|^{2p}}, \tag{13}$$

$$v_{ij}[\mathbf{R}(k)]=\frac{1}{|\mathbf{R}(k)-\mathbf{R}^{ij}|^{2p}}. \tag{14}$$

This weight is large if $\mathbf{R}(k)$ is close to a large number of trajectory points and is far from the points of the existing data set. The configuration $\mathbf{R}(k_{\max})$ with the highest weight is selected, and the index $i_{\max}$ of the nearest IRC point which maximizes $v_{i1}[\mathbf{R}(k_{\max})]$ is identified. The new configuration is added to the set $\{\mathbf{R}^{i_{\max}j}\}$, increasing the number of its elements. This procedure is repeated until the convergence is achieved for the dynamical results from the PES.

III. CALCULATION

The interpolation scheme described above was tested for the following five-atom system,

$$CH_2OH^+ \rightarrow CHO^+ + H_2. \tag{15}$$

An analytic PES for this system was reported by Lee, Park, and Kim[15] (called the LPK surface hereafter). Even though the main goal of the present work is to obtain an interpolated analytic PES from *ab initio* data, the use of the analytic PES makes the test easier because the energy and its derivatives can be calculated at negligible cost. An additional advantage is that the energy partitioning in the products was already

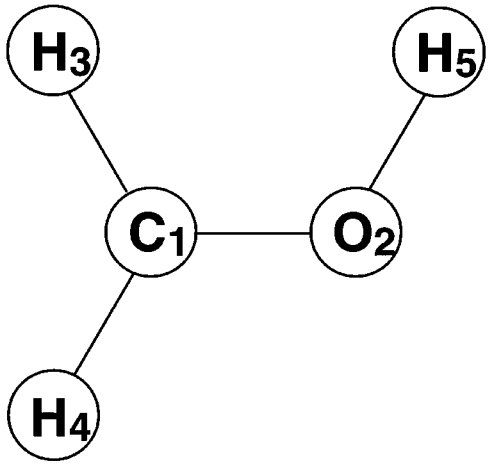

FIG. 1. Atomic numbering for the CH_2OH^+ system.

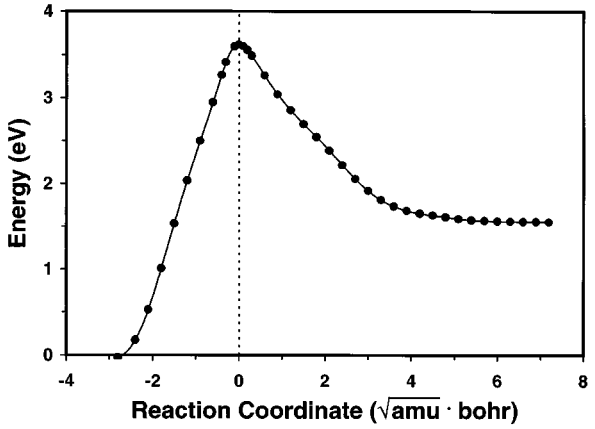

FIG. 2. Potential energy along the recalculated IRC on the interpolated surface obtained with 40 data points on the IRC. The zero of the reaction coordinate corresponds to the transition state. Circles represent the position of the data points.

calculated by the classical trajectory method on the LPK surface. Hence, the convergence of the interpolated surface can be checked easily.

There can be a number of choices as the components of *Z*-matrix type internal coordinates. The interpolated analytic PES using four different sets of coordinates will be compared with one another to see if the selection of coordinates is critical to the success of the overall scheme. These sets are listed in Table I (see Fig. 1). The transformation matrix element J_{ij} and its derivatives in Eqs. (2) and (3) were calculated by adapting some routines in the GAMESS package.[39] The relatively small value of 6 was used for p in Eqs. (9), (11), (13) and (14).

The IRC on the LPK surface was calculated using the constrained optimization method developed by Gonzalez and Schlegel[40,41] with the stride value of 0.1 $\sqrt{\text{amu}}\cdot$ bohr. Out of 101 IRC configurations 38 data points were selected following the method in Sec. II C, within a tolerance of 0.01 eV. In addition, two points in the asymptotic region of the products were added to ensure the proper behavior in this region. Fig. 2 shows the variation of the energy along the IRC of the interpolated surface obtained with these points. The IRC profile of the LPK surface is almost exactly coincidental to Fig. 2. This shows that the potential energy along the IRC can be well represented with a relatively small number of points.

TABLE I. Components of four sets of the *Z*-matrix type internal coordinate system.[a]

	ZT1	ZT2	ZT3	ZT4
Bond length	r_{21}	r_{21}	r_{21}	r_{21}
	r_{31}	r_{31}	r_{41}	r_{31}
	r_{41}	r_{41}	r_{52}	r_{41}
	r_{53}	r_{53}	r_{35}	r_{52}
Bond angle	θ_{312}	θ_{312}	θ_{412}	θ_{312}
	θ_{412}	θ_{413}	θ_{521}	θ_{412}
	θ_{531}	θ_{531}	θ_{352}	θ_{521}
Dihedral angle	ϕ_{3124}	ϕ_{4132}	ϕ_{5214}	ϕ_{3124}
	ϕ_{5312}	ϕ_{5312}	ϕ_{3521}	ϕ_{5214}

[a]Atomic numbering is defined in Fig. 1. The coordinate numbering follows the convention in Chap. 5, Ref. 18.

The classical trajectories were calculated on the above interpolated PES using the same strategy reported for the LPK surface.[15] Briefly, a total of 3000 initial conditions were selected by orthantlike sampling at the transition state structure; the molecular geometry was fixed at the saddle point and the momentum was selected randomly by the techniques developed by Bunker and Hase.[42,43] Integration of the Hamilton's equations of motion was carried out using Gear's variable-order and variable-step algorithm[44] with a step size of 5 fs. The trajectories were integrated until the dissociation was complete (when the two fragments were 5 Å apart) or the system relaxed to the reactant (when the normalized reaction path weight B_i was the largest for the reactant). The energy conservation $|\Delta E|/E$ was within 10^{-5} after completion of the integration. The translational, rotational, and vibrational energies of the fragments were calculated according to the methods in the MERCURY program[45] developed by Nyman, Rynefors, and Hase,[46] and the mode specific effect was calculated according to Park and Bowman.[47]

The configurations on each trajectory were sampled at every 5 fs, collecting up to 20 configurations for each trajectory. Among these, the configurations with the center of mass distance larger than that at the IRC end point (product side) were discarded. Among the configurations collected from a batch of 10 trajectories, the point with the largest weight (Eq. (12)) was selected and added to the data set. Consequently, 300 new data points were added to the initial set after completion of 3000 trajectories.

Convergence of the surface was checked whenever the number of newly added points reached 50. Only 300 initial conditions were used to reduce the calculation cost. Since the trajectory starts from the transition state in the present work, the reaction probability is always around 0.5. Therefore, it cannot be used as a guideline for the convergence check. The

energy partitioning pattern of the products can be used instead. Since most of the excess energy is partitioned into the relative translation of products, or kinetic energy release (KER),[15] the average KER was used for the convergence check. Namely, a full convergence here means that the average KERs obtained by trajectory calculations on the LPK and the interpolated potential energy surfaces are essentially identical.

The trajectory calculations were performed on an IBM 9076 SP2 computer with parallelization of the computer code. Typically, calculation of 3000 trajectories on the interpolated surface constructed with 340 data points took 1.5 h of wall time with 16 CPUs in parallel.

IV. RESULTS AND DISCUSSION

A. Convergence enhancement by the reaction path weight

The influence of the weighting schemes on the convergence is compared in Fig. 3. The change in the average KER with the increase of points in the data set calculated with the weighting scheme in the original IC algorithm is shown in Fig. 3(a). Regardless of the coordinates used, convergence is not satisfactory with this algorithm, which indicates that this scheme is not adequate for the present five-atom case. When the surfaces were generated with the reaction path weight, the average KERs converged to the LPK value well except with the ZT4 coordinate system, as shown in Fig. 3(b). This enhancement of the convergence may be explained in terms of the errors induced by the changes in the transverse curvatures as follows.

Suppose that the potential can be approximated according to the reaction path Hamiltonian formalism,[23,24]

$$V(s,q_1,\cdots,q_{3N-7}) = V_0(s) + \sum_{k=1}^{3N-7} \tfrac{1}{2}\omega_k(s) q_k^2 . \tag{16}$$

The gradient and Hessian with respect to the $3N-6$ dimensional normal coordinates $\mathbf{Q}=(s,q_1,\cdots,q_{3N-7})$ are given by

$$\mathbf{g} = \begin{pmatrix} V_0' + \tfrac{1}{2}\mathbf{q}^T \cdot \boldsymbol{\omega}' \cdot \mathbf{q} \\ \boldsymbol{\omega} \cdot \mathbf{q} \end{pmatrix}, \tag{17}$$

$$\mathbf{H} = \begin{pmatrix} V_0'' + \tfrac{1}{2}\mathbf{q}^T \cdot \boldsymbol{\omega}'' \cdot \mathbf{q} & \mathbf{q}^T \cdot \boldsymbol{\omega}' \\ \boldsymbol{\omega}' \cdot \mathbf{q} & \boldsymbol{\omega} \end{pmatrix}, \tag{18}$$

respectively. Here, $\boldsymbol{\omega}$ denotes the diagonal matrix of the harmonic frequencies of the transverse modes,

$$\boldsymbol{\omega} = \mathrm{diag}(\omega_1,\cdots,\omega_{3N-7}). \tag{19}$$

Let $\mathbf{A}=(s_a,\mathbf{q}_a)$ and $\mathbf{B}=(s_b,\mathbf{0})$ be the two adjacent data points, and $\mathbf{X}=(s_b,\mathbf{q}_b)$ be a trajectory point. The potential at $\mathbf{X}$ dictated by Eq. (16) is

$$V(\mathbf{Q_X}) = V_0(s_b) + \tfrac{1}{2}\mathbf{q}_b{}^T \cdot \boldsymbol{\omega}(s_b) \cdot \mathbf{q}_b . \tag{20}$$

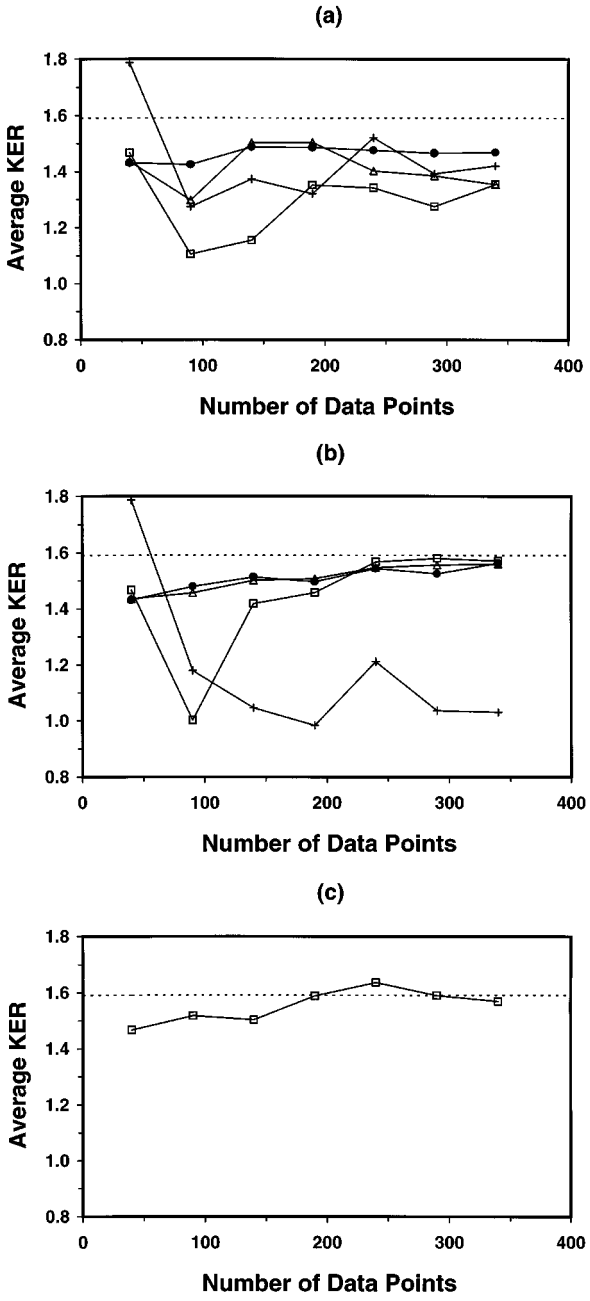

FIG. 3. The average KER vs the number of data points defining the interpolated surface (a) with the original Ichtwan and Collins algorithm and (b) with the present algorithm with ZT1 (△), ZT2 (□), ZT3 (●), and ZT4 (+) coordinate systems. (c) The stabilized result with the ZT2 system using the $N+10$ scheme. Broken lines represent the LPK value.

The same expression is obtained when the potential at $\mathbf{X}$ is given as the Taylor expansion from $\mathbf{B}$, namely $V_{\mathbf{X}\leftarrow\mathbf{B}}$. On the other hand, the same potential may be expressed as the Taylor expansion from $\mathbf{A}$ as follows.

$$V_{\mathbf{X}\leftarrow\mathbf{A}} = V(\mathbf{Q_A}) + (\mathbf{Q_X}-\mathbf{Q_A})^T \cdot \mathbf{g}(\mathbf{Q_A}) + \tfrac{1}{2}(\mathbf{Q_X}-\mathbf{Q_A})^T \cdot \mathbf{H} \cdot (\mathbf{Q_X}-\mathbf{Q_A}). \tag{21}$$

If $\Delta s = s_b - s_a$ is small, the following approximations can be made.

$$V_0(s_b) = V_0(s_a) + V_0'(s_a)\Delta s + \tfrac{1}{2} V_0''(s_a)\Delta s^2, \tag{22}$$

$$\boldsymbol{\omega}(s_b) = \boldsymbol{\omega}(s_a) + \boldsymbol{\omega}'(s_a)\Delta s + \tfrac{1}{2}\boldsymbol{\omega}''(s_a)\Delta s^2. \tag{23}$$

Substituting Eqs. (22) and (23) into Eq. (20), one obtains

$$V_{\mathbf{X}\leftarrow\mathbf{B}} = [V_0(s_a) + V_0'(s_a)\Delta s + \tfrac{1}{2} V_0''(s_a)\Delta s^2] + \tfrac{1}{2}\mathbf{q}_b{}^T \cdot [\boldsymbol{\omega}(s_a) + \boldsymbol{\omega}'(s_a)\Delta s + \tfrac{1}{2}\boldsymbol{\omega}''(s_a)\Delta s^2] \cdot \mathbf{q}_b \tag{24}$$

and

$$V_{\mathbf{X}\leftarrow\mathbf{A}} = V_0(s_a) + \tfrac{1}{2}\mathbf{q}_a{}^T \cdot \boldsymbol{\omega} \cdot \mathbf{q}_a + (V_0'(s_a) + \tfrac{1}{2}\mathbf{q}_a{}^T \cdot \boldsymbol{\omega}' \cdot \mathbf{q}_a)\Delta s + \Delta\mathbf{q}^T \cdot \boldsymbol{\omega} \cdot \mathbf{q}_a + \tfrac{1}{2}(V_0''(s_a) + \tfrac{1}{2}\mathbf{q}_a{}^T \cdot \boldsymbol{\omega}'' \cdot \mathbf{q}_a)\Delta s^2 + \tfrac{1}{2}\Delta\mathbf{q}^T \cdot \boldsymbol{\omega} \cdot \Delta\mathbf{q} + \Delta\mathbf{q}^T \cdot \boldsymbol{\omega}' \cdot \mathbf{q}_a \Delta s. \tag{25}$$

Then, the error in the expansion $V_{\mathbf{X}\leftarrow\mathbf{A}}$ can be approximated as

$$\varepsilon = V_{\mathbf{X}\leftarrow\mathbf{A}} - V_{\mathbf{X}\leftarrow\mathbf{B}} = \tfrac{1}{2}\Delta\mathbf{q}^T \cdot (\boldsymbol{\omega}(s_a) - \boldsymbol{\omega}(s_b)) \cdot \Delta\mathbf{q} - \tfrac{1}{2}\Delta\mathbf{q}^T \cdot \boldsymbol{\omega}''(s_a) \cdot \mathbf{q}_a \Delta s^2, \tag{26}$$

where $\Delta\mathbf{q} = \mathbf{q}_b - \mathbf{q}_a$. Ignoring the second term, ε is simply given by

$$\varepsilon = \tfrac{1}{2}\sum_{k=1}^{3N-7} \Delta\omega_k (q_{b}{}_k - q_{a}{}_k)^2, \tag{27}$$

with

$$\Delta\omega_k = \omega_k(s_a) - \omega_k(s_b). \tag{28}$$

When the transverse curvature changes along the IRC, namely $\Delta\omega_k$ is nonzero, and when there are nonzero $q_{b}{}_k - q_{a}{}_k$ terms, $V_{\mathbf{X}\leftarrow\mathbf{A}}$ and $V_{\mathbf{X}\leftarrow\mathbf{B}}$ will differ ($\varepsilon \neq 0$). Without the reaction path weight, both $V_{\mathbf{X}\leftarrow\mathbf{A}}$ and $V_{\mathbf{X}\leftarrow\mathbf{B}}$ contribute to the potential at **X** if **X** is located halfway between **A** and **B**. This leads to an erroneous value for the potential at **X**. The schematic diagram of such a situation in three dimensional space is illustrated in Fig. 4. An occurrence of this type of error will become more frequent with larger systems, because the number of nonzero terms will increase as the system gets larger. The above difficulty can be mostly removed by using the reaction path weight; $V_{\mathbf{X}\leftarrow\mathbf{A}}$ would have much smaller weight than $V_{\mathbf{X}\leftarrow\mathbf{B}}$ because the reaction path point of **A**, $(s_a, \mathbf{0})$, cannot be closer to **X** than **B** is.

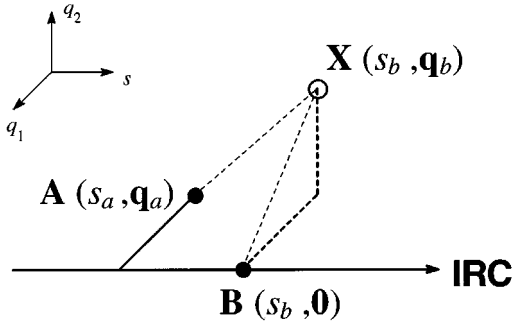

FIG. 4. A schematic representation of the case when a change of the harmonic frequencies of transverse modes may cause an interpolation error. s denotes the reaction coordinate while q_1 and q_2 denote the direction of the normal vectors of transverse modes. **X** is at the same distance from an off-IRC point **A** and an on-IRC point **B**.

B. Characteristics of the interpolated surface

As shown in Figs. 3(a) and 3(b), the results from all of the coordinate types except ZT4 are similar. The anomalous behavior with ZT4 is not unexpected, since it does not contain the H_3–H_5 bond length as its component which becomes important as the reaction proceeds. Therefore, one can conclude that the characteristics of the interpolated surface are rather insensitive to the choice of the coordinates as far as they look chemically adequate to describe the reaction. Hereafter, we will present only the results with the ZT1 coordinate system unless otherwise indicated. It is also interesting to note from Fig. 3(b) that the average KER of 1.43 eV obtained on the surface constructed only with 40 IRC points is not so different from that of 1.60 eV on the LPK surface. This implies that the potential energy surface may be represented reasonably well even for a much larger system where it is difficult to obtain *ab initio* data at many points, as was noted by Jordan, Thompson and Collins.[31]

Fig. 5 compares the contour maps of the LPK surface and the interpolated surface obtained with 340 data points. Except for some minor mismatches, the interpolated surface represents the analytic LPK surface very well. When 3000 trajectories were calculated on these surfaces, 1430 were reactive on the LPK surface while 1451 were reactive on the interpolated one. Fig. 6 compares the kinetic energy release distribution (KERD) obtained by classical trajectory calculation on the interpolated surface with the one on the LPK surface. An excellent agreement between these two KERDs is obvious from the figure. The rotational distributions for the H_2 and CHO^+ products are shown in Figs. 7(a) and 7(b), respectively. A very good agreement is also observed between the distributions calculated on the interpolated and LPK surfaces. The product vibrational distributions also display excellent agreement between the two surfaces, as shown in Figs. 8(a)–8(e). This clearly indicates that the interpolated surface has converged satisfactorily such that the dynamics on it are nearly identical to that on the LPK surface.

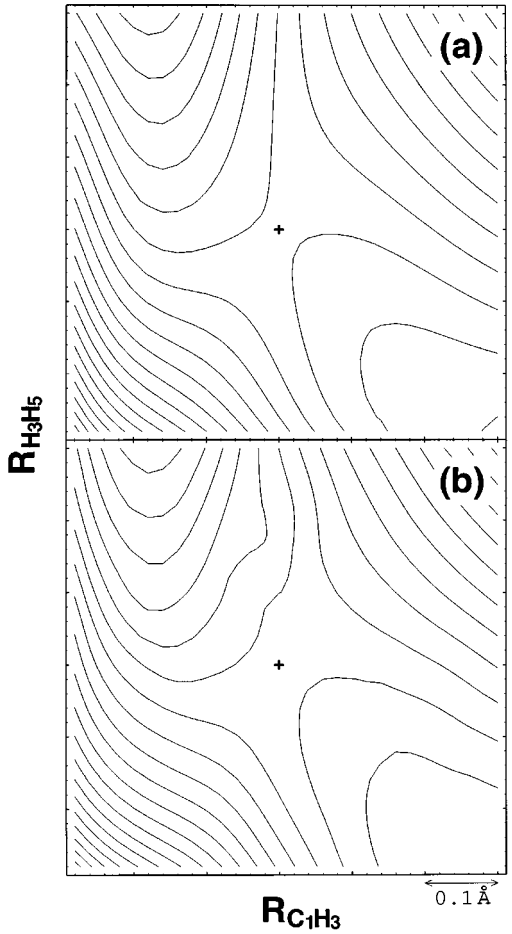

FIG. 5. Contour maps of (a) the LPK and (b) the interpolated surfaces near the transition state. (+) at the center marks the saddle point. Drawn by varying r_{31} and r_{53} with other coordinates fixed to the values at the saddle point.

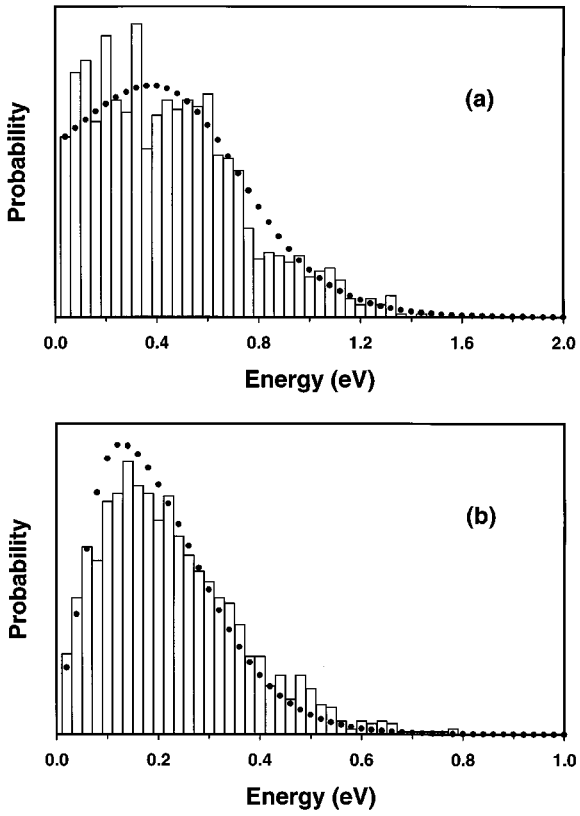

FIG. 7. Rotational energy distributions of the products (a) H_2 and (b) CHO^+. The bars represent the trajectory results on the interpolated surface and the dotted curves represent those on the analytic LPK surface.

C. Stability of the interpolated surface

Because the weighting function of the form in Eqs. (9) and (11) is very sensitive to the variation of the distance, the normalized weighting factor changes sharply at the midpoint of two adjacent configurations. As a result, there will be an abrupt change of the potential if the Taylor expansions from these configurations do not match at the midpoint. A surface with such artifacts is unstable, since the potential derivatives will display severe oscillations. During the course of the data set construction, the anharmonicity correction for transverse modes may result in such kind of instability. When the surface is generated only with points on the IRC, the anharmonicities of transverse modes would lead to some error for any location off the IRC (Fig. 9(a)). This is corrected when new points off the IRC are added to the data set. However, because a point located farthest from the existing data points is generally selected, the surface may develop irregular structures in the course of the improvement as shown in Fig. 9(b). Such irregularities may act as friction to the recoiling of the two fragments. As a consequence, some of the translational energy along the IRC may flow into other degrees of freedom. This may account for the fact that the average KER with ZT2 was lowered at the initial stage in Fig. 3(b). It should be noted that when the average KER on the primitive surface is smaller than the true value (LPK), this lowering may not always occur because the surface is improving anyway with the data set update.

The above difficulty may be remedied sometimes by adding a cluster of configurations at a time. When many

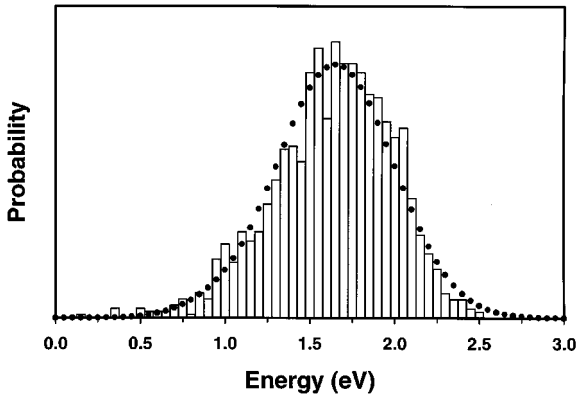

FIG. 6. Kinetic energy release distribution (KERD) of the reaction. The bars represent the classical trajectory result on the interpolated surface and the dotted curve represents that on the analytic LPK surface.

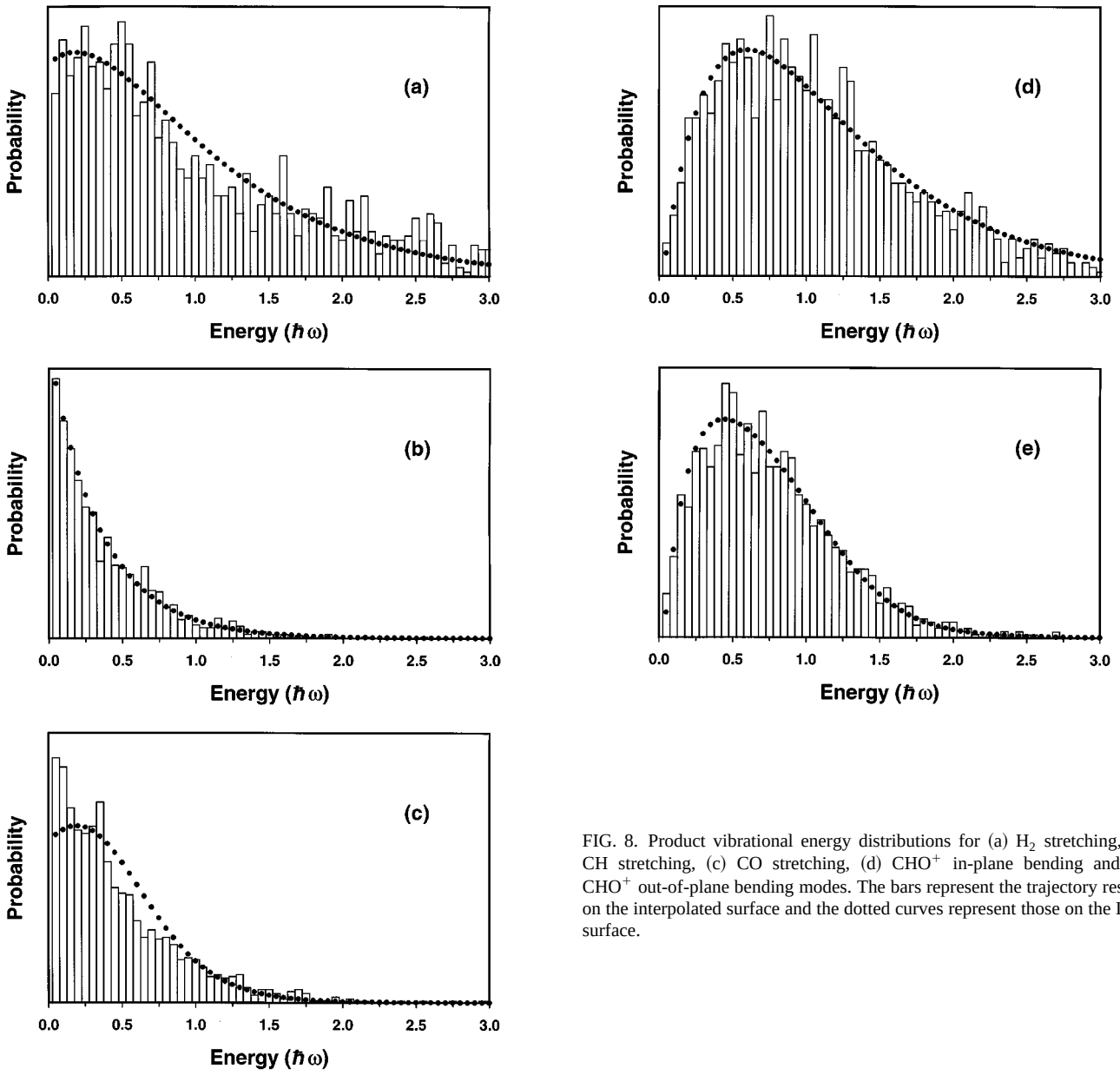

FIG. 8. Product vibrational energy distributions for (a) H_2 stretching, (b) CH stretching, (c) CO stretching, (d) CHO^+ in-plane bending and (e) CHO^+ out-of-plane bending modes. The bars represent the trajectory results on the interpolated surface and the dotted curves represent those on the LPK surface.

points are close in the configuration space, forming a cluster, each point in the cluster will have a high weight because such points have large numerators in Eq. (12). If many points, rather than one, are selected at a time, chances are that points in a cluster will be chosen simultaneously. Let $N+10$ scheme denote the method in which the data set is updated by adding 10 points at a time from a batch of 100 trajectories, while $N+1$ denotes the original method of adding 1 point at a time from a batch of 10 trajectories. When the $N+10$ scheme is used with ZT2, the convergence pattern of the average KER becomes much less oscillatory [Fig. 3(c) versus 3(b)]. A schematic of this situation is illustrated in Fig. 9(c). The surface irregularities are diminished by the addition of a cluster of points, even though the rate of convergence may slow down.

V. SUMMARY AND CONCLUSION

An improved algorithm to construct a potential energy surface of a polyatomic (five or more atoms) system was developed. The main idea was to introduce an additional weighting function associated with the location of a point along the intrinsic reaction coordinate, namely the reaction path weight. Initially, it was thought that the selection of a proper set of internal coordinates may be critical to the successful application of the interpolation scheme to polyatomic systems. Investigation, however, has shown that the dynamical properties of the constructed surface are rather insensitive to such a selection as long as a chemically reasonable set is chosen. This opens the way for the construction of the analytic potential energy surface of large polyatomic systems by

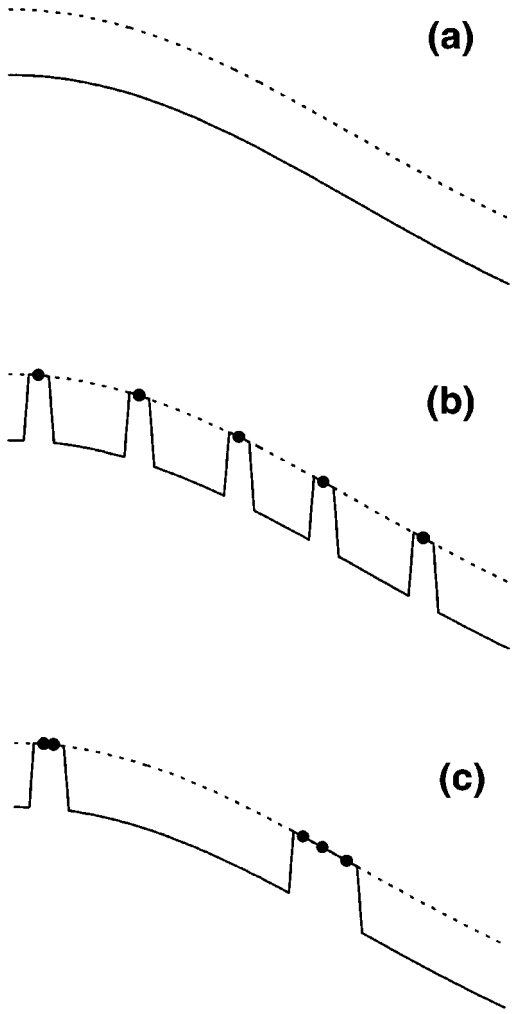

FIG. 9. A schematic representation of the surface irregularities during the data set construction. Broken lines represent the correct surfaces. The solid line in (a) represents the interpolated surface constructed with a data set obtained from on-IRC points. The solid lines in (b) and (c) represent the interpolated surfaces updated with the off-IRC points by the $N+1$ scheme and $N+10$ schemes, respectively. Circles denote the positions of the off-IRC data points.

interpolation of *ab initio* quantum chemical data. The present algorithm was applied to the five-atom reaction system, reaction (15), for which an analytic PES was reported together with the trajectory results for the energy partitioning in the products. The energy partitioning results calculated on the interpolated PES were essentially identical to those on the original analytic PES, demonstrating that the present approach is adequate for five-atom reactions. We are attempting to investigate the applicability of the method to a six or more atom reaction at the moment.

The reaction path Hamiltonian (RPH) formalism proposed by Miller and co-workers[23] also utilizes the concept of the reaction path. Even though the RPH approach attempts to construct the Hamiltonian for a dynamical study, an approximate potential energy surface around the IRC can be obtained with the formalism. However, the potential energy surfaces constructed by the present method would be more accurate because no restriction, such as the harmonic approximation made in RPH for vibrations orthogonal to the reaction coordinate is imposed on the potential shape. The harmonic approximations mentioned above also mean that the potential energy evaluated with RPH may not be adequate when a trajectory probes a region far off the intrinsic reaction path where the effect of anharmonicity can be important. Since the construction of PES by interpolation utilizes information at off-IRC points designated by trajectory results, it is expected that the above problem may be avoided. It is not yet clear, however, whether the present algorithm can deal with the trajectories really far from the IRC because a problem can still arise, for example, in the calculation of the reaction path weight. Such a possibility is also being investigated now.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education, Republic of Korea, under Project No. BSRI-96-3419, and by the Center for Molecular Catalysis and the Korea Science and Engineering Foundation. One author (S.C.P.) thanks the Korea Science and Engineering Foundation and the Ministry of Education of Korea for financial support. A second author (T.G.L.) acknowledges a postdoctoral fellowship from the Research Institute for Basic Science, Seoul National University where he is the Research Fellow. The authors are grateful to the Educational and Research Computing Center at Seoul National University for access to the IBM SP2 computer. The publication cost was supported in part by the Research Institute of Molecular Sciences.

[1] D. G. Truhlar, R. Steckler, and M. S. Gordon, Chem. Rev. **87**, 217 (1987).
[2] I. Shavitt, R. M. Stevens, F. L. Minn, and M. Karplus, J. Chem. Phys. **48**, 2700 (1968).
[3] B. Liu, J. Chem. Phys. **58**, 1925 (1973).
[4] D. G. Truhlar and C. J. Horowitz, J. Chem. Phys. **68**, 2466 (1978); **71**, 1514 (1979).
[5] A. J. C. Varandas, F. B. Brown, C. A. Mead, and D. G. Truhlar, J. Chem. Phys. **86**, 6258 (1987).
[6] R. Schinke and W. A. Lester, Jr., J. Chem. Phys. **70**, 4893 (1979).
[7] J. M. Bowman, J. S. Bittman, and L. B. Harding, J. Chem. Phys. **85**, 911 (1986).
[8] P. R. Bunker, M. Kofranek, H. Lischka, and A. Karpfen, J. Chem. Phys. **89**, 3002 (1988).
[9] C. A. Parish and C. E. Dykstra, J. Chem. Phys. **101**, 7618 (1994).
[10] G. C. Lynch, D. G. Truhlar, F. B. Brown, and J. Zhao, J. Phys. Chem. **99**, 207 (1995).
[11] W. L. Hase, R. J. Wolf, and C. S. Sloane, J. Chem. Phys. **71**, 2911 (1979).
[12] L. M. Raff, J. Phys. Chem. **91**, 3266 (1987).
[13] S. A. Abrash, R. W. Zehner, G. J. Mains, and L. M. Raff, J. Phys. Chem. **99**, 2959 (1995).
[14] B. M. Rice, G. F. Adams, M. Page, and D. L. Thompson, J. Phys. Chem. **99**, 5016 (1995).
[15] T. G. Lee, S. C. Park, and M. S. Kim, J. Chem. Phys. **104**, 4517 (1996).
[16] I. S. Y. Wang and M. Karplus, J. Am. Chem. Soc. **95**, 8160 (1973).
[17] K. K. Baldridge, M. S. Gordon, R. Steckler, and D. G. Truhlar, J. Phys. Chem. **93**, 5107 (1989).
[18] W. J. Hehre, L. Radom, P. v.R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986).
[19] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. Quantum Chem. Symp. **13**, 225 (1979).

[20] T. Helgaker, E. Uggerud, and H. J. Aa. Jensen, Chem. Phys. Lett. **173**, 145 (1990).
[21] E. Uggerud and T. Helgaker, J. Am. Chem. Soc. **114**, 4265 (1992).
[22] W. Chen, W. L. Hase, and H. B. Schlegel, Chem. Phys. Lett. **228**, 436 (1994).
[23] W. H. Miller, N. C. Handy, and J. E. Adams, J. Chem. Phys. **72**, 99 (1980).
[24] W. H. Miller, J. Phys. Chem. **87**, 3811 (1983).
[25] D. G. Truhlar, A. D. Isaacson, and B. C. Garrett, *Theory of Chemical Reaction Dynamics*, edited by M. Baer (Chemical Rubber, Boca Raton, 1985), Vol. IV, Chap. 2.
[26] C. Doubleday, Jr., J. W. McIver, Jr., and M. Page, J. Phys. Chem. **92**, 4367 (1988).
[27] B. C. Garrett, T. Joseph, T. N. Truong, and D. G. Truhlar, Chem. Phys. **136**, 271 (1989).
[28] E. E. Aubanel and D. M. Wardlaw, J. Phys. Chem. **93**, 3117 (1989).
[29] J. Espinosa-García and J. C. Corchado, J. Chem. Phys. **101**, 1333 (1994).
[30] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[31] M. J. T. Jordan, K. C. Thompson, and M. A. Collins, J. Chem. Phys. **102**, 5647 (1995).
[32] M. J. T. Jordan, K. C. Thompson, and M. A. Collins, J. Chem. Phys. **103**, 9669 (1995).
[33] M. J. T. Jordan and M. A. Collins, J. Chem. Phys. **104**, 4600 (1996).
[34] K. A. Nguyen, I. Rossi, and D. G. Truhlar, J. Chem. Phys. **103**, 5522 (1995).
[35] R. A. Marcus, J. Chem. Phys. **45**, 4493 (1966).
[36] D. G. Truhlar and A. Kuppermann, J. Am. Chem. Soc. **93**, 1840 (1971).
[37] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
[38] E. B. Wilson, Jr., J. C. Decius, and P. C. Cross, *Molecular Vibrations* (McGraw-Hill, New York, 1955), Chap. 4.
[39] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[40] C. Gonzalez and H. B. Schlegel, J. Chem. Phys. **90**, 2154 (1989).
[41] C. Gonzalez and H. B. Schlegel, J. Phys. Chem. **94**, 5523 (1990).
[42] D. L. Bunker and W. L. Hase, J. Chem. Phys. **59**, 4621 (1973); W. L. Hase, J. Chem. Phys. **69**, 4711 (1978).
[43] W. L. Hase and D. G. Buckowski, Chem. Phys. Lett. **74**, 284 (1980).
[44] G. Hall and J. M. Watt, *Modern Numerical Methods for Ordinary Differential Equations* (Clarendon, Oxford, 1976).
[45] W. L. Hase, QCPE Program No. 453 **3**, 453 (1983).
[46] G. Nyman, K. Rynefors, and W. L. Hase, Chem. Phys. **110**, 27 (1986).
[47] S. C. Park and J. M. Bowman, Chem. Phys. Lett. **119**, 275 (1985).