RESEARCH ARTICLE | DECEMBER 12 2005

Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method

Andreas Heyden; Alexis T. Bell; Frerich J. Keil

J. Chem. Phys. 123, 224101 (2005)
https://doi.org/10.1063/1.2104507

Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method

Andreas Heyden[a),b)]
Department of Chemical Engineering, Hamburg University of Technology, D-21073 Hamburg, Germany

Alexis T. Bell[a),c)]
Department of Chemical Engineering, University of California, Berkeley, California 94720-1462

Frerich J. Keil
Department of Chemical Engineering, Hamburg University of Technology, D-21073 Hamburg, Germany

(Received 10 May 2005; accepted 12 September 2005; published online 12 December 2005)

A combination of interpolation methods and local saddle-point search algorithms is probably the most efficient way of finding transition states in chemical reactions. Interpolation methods such as the growing-string method and the nudged-elastic band are able to find an approximation to the minimum-energy pathway and thereby provide a good initial guess for a transition state and imaginary mode connecting both reactant and product states. Since interpolation methods employ usually just a small number of configurations and converge slowly close to the minimum-energy pathway, local methods such as partitioned rational function optimization methods using either exact or approximate Hessians or minimum-mode-following methods such as the dimer or the Lanczos method have to be used to converge to the transition state. A modification to the original dimer method proposed by [Henkelman and Jónnson J. Chem. Phys. **111**, 7010 (1999)] is presented, reducing the number of gradient calculations per cycle from six to four gradients or three gradients and one energy, and significantly improves the overall performance of the algorithm on quantum-chemical potential-energy surfaces, where forces are subject to numerical noise. A comparison is made between the dimer methods and the well-established partitioned rational function optimization methods for finding transition states after the use of interpolation methods. Results for 24 different small- to medium-sized chemical reactions covering a wide range of structural types demonstrate that the improved dimer method is an efficient alternative saddle-point search algorithm on medium-sized to large systems and is often even able to find transition states when partitioned rational function optimization methods fail to converge. © *2005 American Institute of Physics*. [DOI: 10.1063/1.2104507]

INTRODUCTION

The harmonic approximation to transition state theory is often used to predict the rates of chemical reactions.[1,2] Such calculations require knowledge of the electronic energy, the geometry, and the real vibrational frequencies at the first-order saddle point (transition state) and at the reactant minimum. While locating potential energy minima can routinely be done, finding saddle points on potential-energy surfaces (PESs) determined from quantum-chemical calculations can be extremely difficult and remains one of the major challenges in chemical kinetics for medium to large systems. Several saddle-point search algorithms have been developed. They can be divided into two groups, those based on interpolation between two minima, and those using only local information. Local surface-walking algorithms explore the PES using local gradient and usually second derivative information. These methods can be initiated anywhere on the PES and when close to a saddle point converge rapidly. Unfortunately, surface-walking algorithms perform poorly for systems with several low-frequency vibrational modes or for searches started far from a transition state. Therefore, one of the major challenges in the use of surface-walking algorithms for large systems is to find a good initial guess of the transition state. An intuitive guess of the transition state configuration often results in a second derivative matrix that has several unstable modes, none of which resemble the motion along the reaction coordinate. Furthermore, even if a transition state is found it is possible that the transition state does not connect reactant and product states.

The most common surface-walking algorithms are quasi-Newton methods introduced by Cerjan and Miller[3] and later modified by Simons and co-workers[4–6] and Wales.[7] One of these methods widely used for *ab initio* molecular calculations is the partitioned rational function optimization algorithm of Baker,[8] which utilizes either an approximate or an exact Hessian matrix. For large systems with a starting geometry far from the saddle point exact Hessians have to be calculated regularly even if a Hessian matrix updating

[a)]Authors to whom correspondence should be addressed.
[b)]Electronic mail: a.heyden@tuhh.de
[c)]Electronic mail: bell@cchem.berkeley.edu

scheme[9–11] is used. If the Hessian matrix cannot be determined analytically, the second derivative matrix has to be determined numerically. For large systems this step often becomes prohibitively expensive.

Recently, minimum-mode-finding algorithms, such as the dimer[12] and Lanczos methods,[13] have been developed. Instead of calculating the full Hessian matrix these algorithms calculate only the lowest eigenvalue and the corresponding eigenvector. The dimer method, developed by Henkelman and Jónsson,[12] was designed to determine which activated transitions can occur from a given initial state at a finite temperature. In this way the algorithm is intended to find a set of low-lying saddle points at the boundary of the potential-energy basin associated with the initial state.

Interpolation methods generate a sequence of configurations located between the reactant and product states. These algorithms convert a saddle-point search in configuration space to a minimization problem in discretized path space. The solution of a fully converged interpolation algorithm in mass-weighted coordinates is the minimum-energy pathway. In this way, interpolation methods guarantee that the saddle point found connects both reactant and product states. Furthermore, unlike saddle-point searches, minimization problems can easily handle large numbers of low-frequency modes. Since the dimensionality of the discretized path space is significantly larger than that of the configurational space the computational cost of interpolation methods per iteration is significantly larger than for most surface-walking algorithms. Henkelman and Jónsson[14] have suggested that an interpolation scheme should only be used to generate a good initial guess for the saddle point and the minimum-energy pathway. This guess is then refined using a local surface-walking algorithm. In this way, a good guess for the saddle point and the imaginary mode connecting both reactant and product states is generated from the interpolation algorithm, and the saddle point found by the surface-walking algorithm is most likely a transition state connecting reactant and product states. A frequently used interpolation algorithm is the nudged-elastic band (NEB).[14,15] While highly effective for many problems, this method initiates the search for a transition state by assuming that the minimum-energy pathway (MEP) is a straight line in multidimensional space connecting the reactant and product states. Peters *et al.*[16] have shown that the growing-string method, an interpolation method that does not require an initial guess for the initial pathway, needs significantly fewer gradient calculations to find a saddle point than the zero-temperature string method of Ren and coworkers[17,18] and the NEB.[14,15]

The purpose of the present study is to identify the factors affecting the performance of the dimer method and to develop an improved version of this method. The improved dimer method is then compared with the well-established partitioned rational function optimization methods for finding transition states. The initial guess for the position of the transition state in Cartesian coordinates was obtained using the growing-string method. The performance of different methods for refining the initial guess for the transition state was tested for 24 different chemical reactions involving small- to medium-sized reactants and covering a wide range of structural types.

THEORY

Quantum-chemical calculations

Quantum-chemical calculations were performed using nonlocal, gradient-corrected density-functional theory[19] (DFT) in C_1 symmetry. The effects of exchange and correlation were described by Becke's 3-parameter hybrid exchange-correlation functional, B3LYP.[20] Unless otherwise stated a self-consistent-field (SCF) convergence criterion of 10^{-7} hartree and the standard grid ($m3$) for numerical quadrature are used.[21] The triple-ζ basis set with polarization functions (TZVP) was used for all atoms to mimic a realistic level of theory.[22] All energy and gradient calculations were carried out using the TURBOMOLE V5.7 suite of programs.[23,24] Convergence was considered to have been achieved for the standard defaults of 3×10^{-4} hartree/bohr on the maximum gradient component and either an energy change from the previous cycle of less than 10^{-6} hartree or a maximum predicted displacement of less than 3×10^{-4} bohr per coordinate.[25]

Growing-string method

The growing-string method developed by Peters *et al.*[16] was used to find approximations to the minimum-energy pathway and the transition state. This method is based on the "zero-temperature string method" of Ren and coworkers[17,18] for finding minimum-energy pathways and involves two steps. In the first, evolution, step a string of nodes is moved so as to minimize the objective function F:

$$F(\varphi(\sigma_0),\varphi(\sigma_1),\ldots,\varphi(\sigma_n))=\sum_{k=0}^{n}|\mathbf{p}(\varphi(\sigma_k))|^2. \quad (1)$$

Here $\mathbf{p}(\varphi(\sigma_k))$ is the gradient of the potential energy in mass-weighted coordinates of configuration $\varphi(\sigma_k)$ from which the tangential component has been projected out. The global minimum of the objective function F is the minimum-energy pathway. In the second, parametrization, step the nodes are redistributed along the string according to a chosen parametrization density. In the growing-string method an initially discontinuous parametrization density is chosen, so that the string grows from its end points until eventually the two ends join to form one continuous string. In this way the number of nodes on the string n changes between iterations. For the present study, we used a uniform distribution of nodes at both ends, starting with two nodes on each side. Either 9 or 13 nodes were used for the fully converged string. Thirteen nodes were used only if none of the local saddle-point search algorithms used subsequently was able to locate the correct transition state starting from an initial configuration generated using a string with 9 nodes. A new node was added to a string fragment (i.e., the parametrization density evolves), if the force on the node at the end of a string fragment, from which the tangential component is projected out, is smaller than 0.02 E_h/bohr. In this way, the string fragments of the

growing-string method merge relatively rapidly and an approximation of the MEP is generated. Once the parametrization density becomes continuous, the string is not converged further to the MEP, the algorithm is stopped, and a local saddle-point search algorithm is started. The initial guess for the transition state is the geometry of the node with the highest energy and the initial path guess is the path tangent at the highest-energy node. It is noted that this procedure does not lead to the exact MEP; however, the MEP can be determined from the converged saddle point by following the gradient down to the neighboring minima.[26–28] Converging the string to the MEP with high accuracy is unnecessary if only saddle points which connect reactant and product states are needed.

Partitioned rational function optimization method

Surface-walking algorithms for finding stationary points are usually based on a local quadratic approximation of the potential-energy surface. The fastest step $\Delta\mathbf{x}$ to reach a stationary point is then the Newton-Raphson (NR) step:

$$\Delta\mathbf{x} = \sum_i (\mathbf{u}_i^T\mathbf{f})\mathbf{u}_i/\lambda_i. \tag{2}$$

Here $\mathbf{u}_i$ is the eigenvector, λ_i is the eigenvalue of the second derivative matrix of the energy E, and $\mathbf{f}$ is the force at the current configuration. Cerjan and Miller[3] introduced a shift parameter γ to the Hessian eigenvalues λ_i to ensure that the surface-walking algorithm moves uphill along the lowest Hessian eigenvector and downhill along all other eigenvectors.

$$\Delta\mathbf{x} = \sum_i (\mathbf{u}_i^T\mathbf{f})\mathbf{u}_i/(\lambda_i - \gamma). \tag{3}$$

Banerjee *et al.*[5] showed that the above equation can be obtained when a rational function optimization (RFO) approach is used to approximate the potential-energy surface. The shift parameter γ can then be found iteratively from the relationship

$$\sum_i (\mathbf{u}_i^T\mathbf{f})^2/(\gamma - \lambda_i) = \gamma. \tag{4}$$

For finding first-order transition states Banerjee *et al.*[5] suggested separating the principle mode to be maximized and the $(n-1)$ modes to be minimized by partitioning the RFO matrix into two smaller partitioned rational function optimization (P-RFO) matrices and solving each problem separately. Hence there are two shift parameters, γ_p and γ_n. For maximization along the kth mode, γ_p is the largest eigenvalue of

$$\begin{pmatrix} \lambda_k & -\mathbf{u}_k^T\mathbf{f} \\ -\mathbf{u}_k^T\mathbf{f} & 0 \end{pmatrix}\begin{pmatrix} \Delta x_k \\ 1 \end{pmatrix} = \gamma_p\begin{pmatrix} \Delta x_k \\ 1 \end{pmatrix} \tag{5}$$

and γ_n is the lowest eigenvalue of

$$\begin{pmatrix} \lambda_1 & & & & & & -\mathbf{u}_1^T\mathbf{f} \\ & \ddots & & & 0 & & \vdots \\ & & \lambda_{k-1} & & & & -\mathbf{u}_{k-1}^T\mathbf{f} \\ & & & \lambda_{k+1} & & & -\mathbf{u}_{k+1}^T\mathbf{f} \\ & 0 & & & \ddots & & \vdots \\ & & & & & \lambda_n & -\mathbf{u}_n^T\mathbf{f} \\ -\mathbf{u}_1^T\mathbf{f} & \cdots & -\mathbf{u}_{k-1}^T\mathbf{f} & -\mathbf{u}_{k+1}^T\mathbf{f} & \cdots & -\mathbf{u}_n^T\mathbf{f} & 0 \end{pmatrix} \times \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_{k-1} \\ \Delta x_{k+1} \\ \vdots \\ \Delta x_n \\ 1 \end{pmatrix} = \gamma_n \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_{k-1} \\ \Delta x_{k+1} \\ \vdots \\ \Delta x_n \\ 1 \end{pmatrix}. \tag{6}$$

The step $\Delta\mathbf{x}$ then has the form

$$\Delta\mathbf{x} = \sum_{j=1}^{n} \Delta\mathbf{x}_j, \quad \Delta\mathbf{x}_k = \mathbf{u}_k^T\mathbf{f}\mathbf{u}_k/(\lambda_k - \gamma_p), \tag{7}$$

$$\Delta\mathbf{x}_{i,i\neq k} = \mathbf{u}_i^T\mathbf{f}\mathbf{u}_i/(\lambda_i - \gamma_n).$$

The mode k is generally chosen to correspond to the lowest eigenvalue, whereas in a mode-following procedure, the mode k followed is the one having the greatest overlap with the mode followed in the previous cycle or the initial path guess from the growing-string method. Care is taken to always ensure that the total step length does not exceed $\Delta x_{\max} = 0.1$ bohr.

Recalculation of the Hessian matrix for every cycle is very time consuming, and hence, for larger systems the second derivative matrix is calculated at predefined intervals and Hessian updating schemes are used between these intervals. The Bofill[11] Hessian updating scheme, $\mathbf{H}_{j+1} = \mathbf{H}_j + \Delta\mathbf{H}_j$, was used in the present study, which combines the updating scheme of Powell[10] with the symmetric, rank-one (SR1) updating scheme of Murtagh and Sargent.[9]

$$\Delta\mathbf{H}_j^{\text{Bofill}} = \phi^{\text{Bofill}}\Delta\mathbf{H}_j^{\text{SR1}} + (1-\phi^{\text{Bofill}})\Delta\mathbf{H}_j^{\text{Powell}}$$

$$\Delta\mathbf{H}_j^{\text{SR1}} = \frac{(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T}{(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T\Delta\mathbf{x}_j}$$

$$\Delta\mathbf{H}_j^{\text{Powell}} = \frac{(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)\Delta\mathbf{x}_j^T + \Delta\mathbf{x}_j(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T}{\Delta\mathbf{x}_j^T\Delta\mathbf{x}_j} - \frac{(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T\Delta\mathbf{x}_j(\Delta\mathbf{x}_j\Delta\mathbf{x}_j^T)}{(\Delta\mathbf{x}_j^T\Delta\mathbf{x}_j)^2} \tag{8}$$

with $\Delta\mathbf{f}_{j+1} = \mathbf{f}(\mathbf{x}_j) - \mathbf{f}(\mathbf{x}_{j+1})$, and the Bofill factor given by

$$\phi^{\text{Bofill}} = \frac{[(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T\Delta\mathbf{x}_j]^2}{(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)^T(\Delta\mathbf{f}_{j+1} - \mathbf{H}_j\Delta\mathbf{x}_j)\Delta\mathbf{x}_j^T\Delta\mathbf{x}_j}. \tag{9}$$

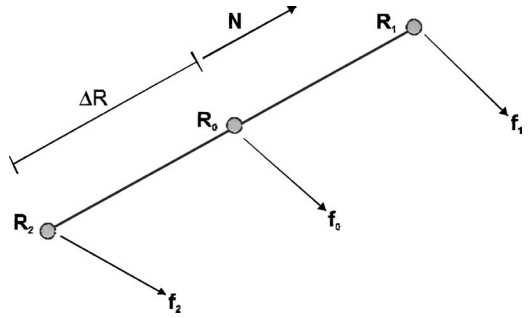

FIG. 1. Definition of the various points in $3n$ dimensional space used by the improved dimer method.

Dimer method

The original dimer method, developed by Henkelman and Jónsson,[12] is a local saddle-point search algorithm that uses only first derivatives of the potential energy and is, therefore, especially appropriate for large systems. The dimer method uses two points in $3n$ dimensional space, which are slightly displaced by a fixed distance, $2\Delta R$ (see Fig. 1). The term n represents the number of atoms. The locations of the dimer end points $\mathbf{R}_1$ and $\mathbf{R}_2$ are defined by

$$\mathbf{R}_1 = \mathbf{R}_0 + \Delta R\mathbf{N} \text{ and } \mathbf{R}_2 = \mathbf{R}_0 - \Delta R\mathbf{N}, \tag{10}$$

where $\mathbf{N}$ is a unit vector along the dimer axis and $\mathbf{R}_0$ is the dimer midpoint. The saddle-point search algorithm involves two steps. In the first step the dimer axis is rotated into the lowest curvature mode of the potential energy at the midpoint of the dimer $\mathbf{R}_0$, and then the dimer is translated for a certain step length on the potential-energy surface thereby moving it towards a saddle point. In significant contrast to quasi-Newton-Raphson algorithms, where the Hessian $\mathbf{H}$ is updated, the curvature of the PES C_x is calculated numerically along one direction (dimer axis) and therefore is more accurate than a curvature calculated with an updated Hessian.

$$C_N = \mathbf{N}^T\mathbf{H}\mathbf{N} \approx \frac{(\mathbf{f}_2 - \mathbf{f}_1)^T \cdot \mathbf{N}}{2\Delta R} \approx \frac{E_1 + E_2 - 2E_0}{\Delta R^2}. \tag{11}$$

Here $\mathbf{f}_1$ and $\mathbf{f}_2$ are the forces acting on configurations 1 and 2 and E_i is the energy at the corresponding configuration. From Eq. (11) it can be seen that during a rotation of the dimer axis (the first step of the algorithm), the minimum of the curvature in the rotation plane is equivalent to the minimum of the dimer energy, $E=E_1+E_2$ (assuming the dimer axis is small enough that the energy varies quadratically). The rotational plane is spanned by the dimer axis and a normalized vector Θ (to be determined) that is orthogonal to the dimer axis. The optimal direction for Θ is that which leads to a maximum overlap (scalar product) with the eigenvector corresponding to the lowest eigenvalue. The steepest descent direction for rotation can be calculated as

$$\frac{\partial E}{\partial \varphi} = \frac{\partial E_1}{\partial \varphi} + \frac{\partial E_2}{\partial \varphi} = -\mathbf{f}_1^T\frac{\partial \mathbf{R}_1}{\partial \varphi} - \mathbf{f}_2^T\frac{\partial \mathbf{R}_2}{\partial \varphi} \tag{12}$$

with

$$\mathbf{R}_{1/2}(\varphi) = \mathbf{R}_0 \pm \Delta R(\mathbf{N}\cos\varphi + \Theta\sin\varphi) \tag{13}$$

and

$$\frac{\partial \mathbf{R}_{1/2}}{\partial \varphi} = \pm \Delta R\Theta \tag{14}$$

and consequently

$$\frac{\partial E}{\partial \varphi} = -\Delta R(\mathbf{f}_1 - \mathbf{f}_2)^T\Theta. \tag{15}$$

It is easily seen from Eq. (15) that the steepest descent direction Θ has to be parallel to $(\mathbf{f}_1-\mathbf{f}_2)$. Θ is then orthogonalized to the dimer axis $\mathbf{N}$. An improvement to the steepest descent direction can be obtained with a modified conjugate gradient approach as described by Henkelman and Jónsson.[12] The efficiency of the dimer algorithm relies heavily on an efficient algorithm for rotating the dimer into the lowest curvature mode spanned by $\mathbf{N}$ and Θ. Henkelman and Jónsson[12] showed that within a second-order Taylor-series expansion of the energy the rotational force F on the dimer can be expressed as

$$-\frac{\partial E}{\partial \varphi} = F \approx A\sin(2\varphi), \tag{16}$$

where A is a constant. Then the lowest curvature mode is found by rotating the dimer axis by

$$\varphi_{\min} = -\frac{1}{2}\arctan\left(\frac{2F}{F'}\right) - \delta\varphi/2, \tag{17}$$

where the rotational force F and the rotational curvature F' are found by doing gradient calculations at a dimer configuration rotated by a small angle $\delta\varphi$. One obtains

$$F = \frac{[(\mathbf{f}_1 - \mathbf{f}_2)^T \cdot \Theta]_{\phi=\delta\varphi} + [(\mathbf{f}_1 - \mathbf{f}_2)^T \cdot \Theta]_{\phi=0}}{2}, \tag{18}$$

$$F' = \frac{[(\mathbf{f}_1 - \mathbf{f}_2)^T \cdot \Theta]_{\phi=\delta\varphi} - [(\mathbf{f}_1 - \mathbf{f}_2)^T \cdot \Theta]_{\phi=0}}{\delta\varphi}. \tag{19}$$

Olsen *et al.*[29] suggested that after rotation if the rotational force F does not lie below a chosen threshold δF, the rotational procedure is repeated multiple times until convergence is achieved or a maximum number of allowed rotations is reached. Thus, at least four gradient calculations are necessary for every rotation.

In the second step of the dimer method, the dimer midpoint is translated along a modified force $\mathbf{f}^\dagger$, where the force component along the dimer axis is inverted. In convex regions of the PES, where the curvature along the dimer axis is positive, the modified force is simply the negative of the force component along the dimer axis. This procedure forces the dimer to leave the convex region faster.

$$\mathbf{f}^\dagger = \begin{cases} -(\mathbf{f}_0^T \cdot \mathbf{N})\mathbf{N} & \text{if } C_N > 0 \\ \mathbf{f}_0 - 2(\mathbf{f}_0^T \cdot \mathbf{N})\mathbf{N} & \text{if } C_N < 0. \end{cases} \tag{20}$$

In the original dimer method the force at the midpoint is calculated as the arithmetic mean of the force at positions 1 and 2. Next, the length of the step is determined by doing a Newton step with a maximum step length $\Delta x_{\max}$.

$$\Delta x = -\frac{F^{\dagger}}{C^{\dagger}} = -\frac{(\mathbf{f}^{\dagger}|_{x=\delta x} + \mathbf{f}^{\dagger}|_{x=0})^T \cdot \mathbf{N}^{\dagger}/2}{(\mathbf{f}^{\dagger}|_{x=\delta x} - \mathbf{f}^{\dagger}|_{x=0})^T \cdot \mathbf{N}^{\dagger}/\delta x} + \frac{\delta x}{2}. \tag{21}$$

Here the curvature along the modified force direction $\mathbf{N}^{\dagger} = \mathbf{f}^{\dagger}/\|\mathbf{f}^{\dagger}\|$ is calculated by first doing a small step with step length δx, calculating gradients, and then finite differencing. A conjugate gradient approach for determining the search direction for dimer translation has been suggested in order to improve the performance of the dimer method.[12] To conclude, two gradient calculations have to be done for every dimer translation, and, therefore, altogether every dimer cycle involves at least six gradient calculations.

The original algorithm of Henkelman and Jónsson[12] was developed for use with an analytical potential-energy surface, for which energies and gradients can be calculated with double precision accuracy. Quantum-chemical SCF calculations cannot be done efficiently with the same accuracy. As a result, small dimer separations of $\Delta R = 10^{-3}$ Å and small displacements during rotations (10^{-3} rad) and translations (10^{-3} Å) as suggested by Henkelman and Jónsson[12] should not be used because of the numerical noise in quantum-chemical potential-energy surfaces. At the same time molecular coordinates have to be converged to 3×10^{-4} bohr, and consequently an algorithm in which the forces are not calculated at the dimer midpoint does not converge efficiently to the saddle point with chemical precision. A modification to the original dimer method has recently been suggested by Heyden and Keil[30] and Olsen *et al.*[29] Instead of calculating the gradient at $\mathbf{R}_2$ the gradient is calculated at $\mathbf{R}_0$ and the force at $\mathbf{R}_2$ is approximated by $\mathbf{f}_2 = 2\mathbf{f}_0 - \mathbf{f}_1$ (linear extrapolation). This decreases the accuracy of the curvature calculation from $O(\Delta R^2)$ to $O(\Delta R)$ but also decreases the minimum number of gradient calculations for rotation to two (twice at $\mathbf{R}_1$) and translation to two (twice at $\mathbf{R}_0$). In addition, the energy and gradient are calculated in this algorithm at the dimer midpoint, so that convergence to a saddle point can be achieved with a significant dimer separation ΔR.

A further improvement to the dimer method, which is well suited for potential-energy surfaces determined from quantum-chemical calculations, has been proposed.[31] A flow sheet for this algorithm is shown in Fig. 2. Instead of working with two points in configuration space as in the original dimer method, the algorithm utilizes the position of the dimer midpoint $\mathbf{R}_0$ and, depending on whether the curvature along the dimer axis is calculated by central difference or forward difference, energies and gradients are calculated at $\mathbf{R}_1$ and $\mathbf{R}_2$. The distance between $\mathbf{R}_1$ and $\mathbf{R}_2$ (central differences) or $\mathbf{R}_1$ and $\mathbf{R}_0$ (forward differences) is chosen to be as small as possible in order to increase the accuracy of the calculated curvature, Eq. (11), and at the same time large enough to avoid numerical noise. For standard quantum-chemical SCF accuracies a value of $\Delta R = 10^{-2}$ bohr is found to be appropriate. If the true Hessian matrix is known, the curvature along any direction $\mathbf{x}$ during rotation is given by

$$C_x = \mathbf{x}^T \mathbf{H} \mathbf{x}. \tag{22}$$

Expanding the direction vector $\mathbf{x}$ in an orthonormal eigenvector basis of the Hessian, $\mathbf{u}_i$, and using

$$\mathbf{x} = \mathbf{N} \cos\varphi + \Theta \sin\varphi, \tag{23}$$

it follows that the curvature during a rotation can always be expressed exactly by a short Fourier series:

$$\begin{aligned} C_x = {} & \cos^2\varphi \left(\sum_i \lambda_i (\mathbf{N}^T \mathbf{u}_i)^2 \right) \\ & + 2 \sin\varphi \cos\varphi \left(\sum_i \lambda_i (\mathbf{N}^T \mathbf{u}_i)(\Theta^T \mathbf{u}_i) \right) \\ & + \sin^2\varphi \left(\sum_i \lambda_i (\Theta^T \mathbf{u}_i)^2 \right) = \frac{a_0}{2} + a_1 \cos(2\varphi) \\ & + b_1 \sin(2\varphi). \end{aligned} \tag{24}$$

Here λ_i is the eigenvalue corresponding to the eigenvector $\mathbf{u}_i$ of the exact Hessian and a_0, a_1, and b_1 are constants that are determined by the eigenvalues and eigenvectors of the Hessian. The minimum curvature in the plane spanned by $\mathbf{N}$ and Θ can therefore be found exactly by rotating the dimer by

$$\varphi_{\min} = \frac{1}{2} \arctan\left(\frac{b_1}{a_1} \right). \tag{25}$$

Substitution of $\varphi_{\min}$ into Eq. (24) gives the minimum curvature without the need for additional energy or gradient calculations. If the curvature after rotation is larger than before rotation a maximum curvature is found and the rotation angle has to be increased by 90°.

The coefficients a_0, a_1, and b_1 appearing in Eqs. (24) and (25) are calculated as accurately and as efficiently as possible. It follows from Eq. (11) that the total dimer energy E and the rotational force $-\partial E/\partial\varphi$ vary approximately during rotation like a Fourier series:

$$E \approx 2E_0 + \Delta R^2 \left(\frac{a_0}{2} + a_1 \cos(2\varphi) + b_1 \sin(2\varphi) \right), \tag{26}$$

$$\begin{aligned} -\frac{\partial E}{\partial \varphi} = \Delta R (\mathbf{f}_1 - \mathbf{f}_2)^T \cdot \Theta \approx \Delta R^2 (2a_1 \sin(2\varphi) \\ - 2b_1 \cos(2\varphi)) = -\Delta R^2 \frac{\partial C_x}{\partial \varphi}. \end{aligned} \tag{27}$$

Here a_0, a_1, and b_1 are the same coefficients as those appearing in Eq. (24). Before rotation, the curvature along $\varphi = 0$ and the derivative of the curvature with respect to the angle φ is known, so that just one more piece of information is needed in order to determine the curvature along any direction in the rotational plane spanned by $\mathbf{N}$ and Θ. One way is to rotate the dimer by a fixed rotation angle φ, e.g., 45°, and then calculate the curvature (one energy calculation if forward differences are used) or the derivative of the curvature (one force calculation if forward differences are used), Eqs. (28) and (29).

$$C_x(\varphi_1) \approx \frac{2(E_1(\varphi_1) - E_0 + \Delta R \mathbf{f}_0^T \cdot \mathbf{N}(\varphi_1))}{\Delta R^2}, \tag{28}$$

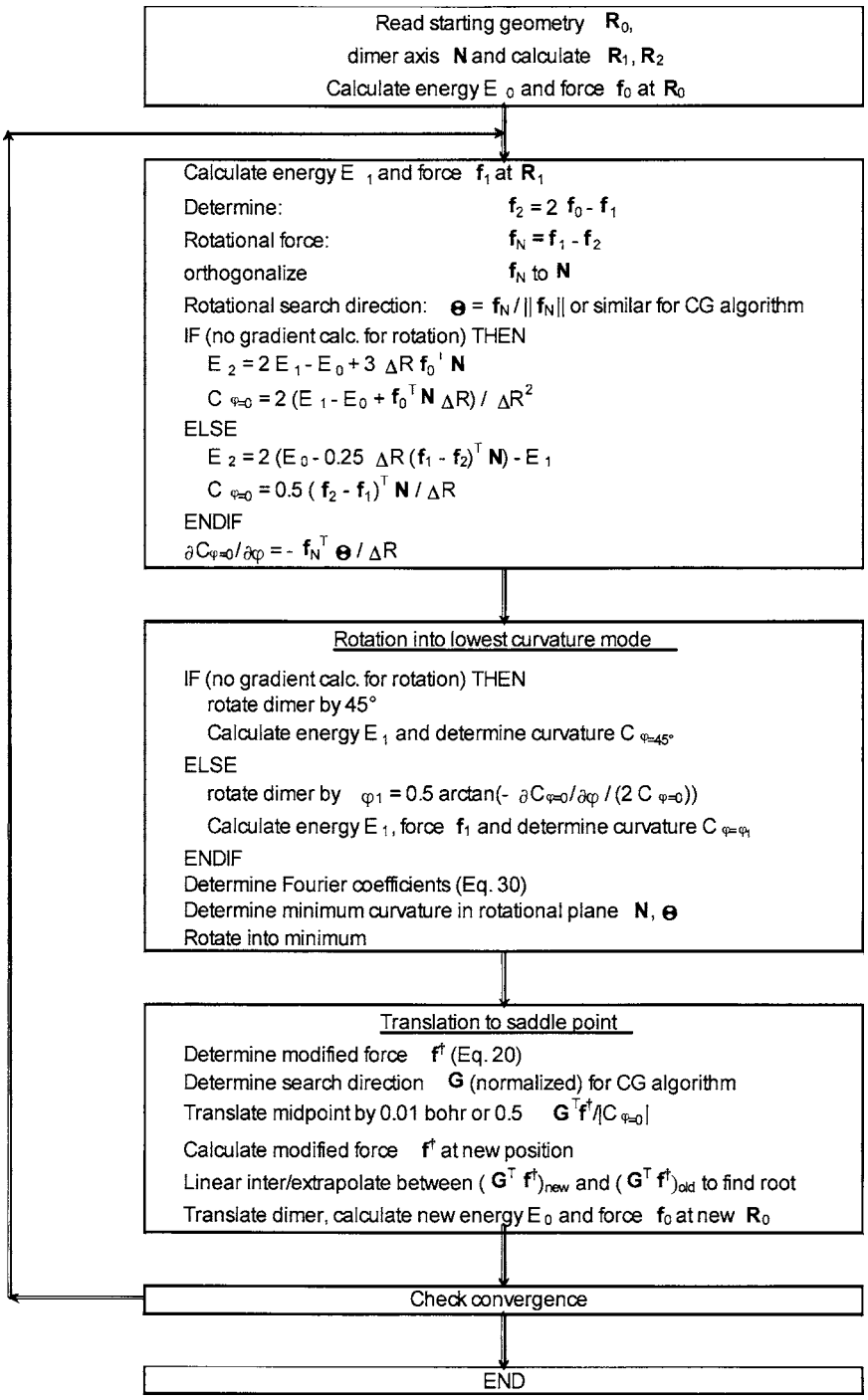

FIG. 2. Flow sheet of the improved dimer method.

$$\left.\frac{\partial C_x}{\partial\varphi}\right|_{\varphi=\varphi_1} \approx \frac{1}{\Delta R}(\mathbf{f}_2(\varphi_1) - \mathbf{f}_1(\varphi_1))^T \cdot \Theta(\varphi_1). \tag{29}$$

In the case of forward differentiation $\mathbf{f}_2$ is again approximated as $\mathbf{f}_2 = 2\mathbf{f}_0 - \mathbf{f}_1$. Thus, ultimately, a_0, a_1, and a_2 are given by

$$a_1 = \frac{C_x|_{\varphi=0} - C_x|_{\varphi=\varphi_1} + 0.5\,(\partial C_x/\partial\varphi)|_{\varphi=0}\sin(2\varphi_1)}{1-\cos(2\varphi_1)}$$

or

$$= \frac{\left.\frac{\partial C_x}{\partial\varphi}\right|_{\varphi=0}\cos(2\varphi_1) - \left.\frac{\partial C_x}{\partial\varphi}\right|_{\varphi=\varphi_1}}{2\sin(2\varphi_1)}$$

$$b_1 = 0.5\left.\frac{\partial C_x}{\partial\varphi}\right|_{\varphi=0},$$

$$a_0 = 2(C_x|_{\varphi=0} - a_1). \tag{30}$$

It is important to note here that due to the numerical noise associated with any quantum-chemical SCF calculation it is not recommended to take a small step and then extrapolate the curvature as suggested by Henkelman and Jónsson[12]

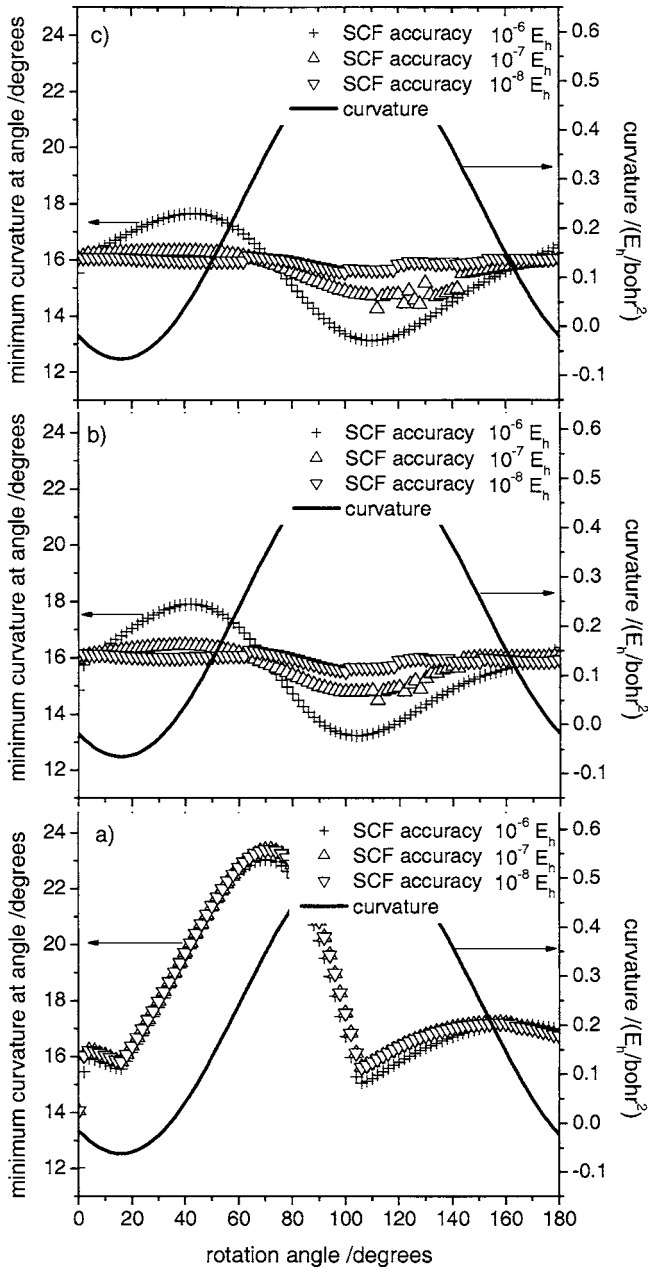

FIG. 3. Effects of initial dimer orientation on the curvature and predicted rotation angle required to achieve minimum curvature in the rotation plane for the isomerization reaction HCN↔HNC. (a) Original dimer method; rotation angle of minimum curvature predicted with Eq. (17). $\delta\phi = 0.01$ rad, $\Delta R = 0.01$ bohr. (b) Improved dimer method with gradient calculation for rotation. (c) Improved dimer method with energy calculation for rotation. In (b) and (c) the rotation angle of minimum curvature was predicted with Eq. (25), $\Delta R = 0.01$ bohr; initial fixed rotation angle of 45°. Forward differences were used in all cases. For a SCF accuracy of 10^{-8} E_h a very fine numerical quadrature grid ($m5$) was used for the DFT calculation.

and Olsen *et al.*,[29] but, rather, to undertake a large rotational step and then interpolate. Figure 3 illustrates the effects of initial dimer orientation on the curvature and predicts rotation angle required to achieve minimum curvature in the rotation plane for the isomerization reaction HCN↔HNC. Depending on the initial direction of the dimer axis the original dimer method predicts a minimum curvature at a rotation angle between 12° and 24° almost independent of the accuracy used in the SCF calculation. Using a SCF accuracy of 10^{-6} E_h the improved dimer method presented here predicts the minimum curvature at a rotation angle between 13° and 18°. This range of predicted rotation angles required to achieve minimum curvature can be reduced significantly by raising the accuracy of the SCF calculation. Using the algorithm proposed here, the dimer axis is always very close to the minimum curvature mode after rotation and the curvature after rotation is determined by Eq. (24). In order to increase the performance of dimer rotation somewhat, the rotational curvature before rotation is approximated and used to determine the initial rotation angle φ_1:

$$\begin{aligned}\frac{\partial^2 C_x}{\partial\varphi^2}\Delta R^2 &\approx \frac{\partial^2 E}{\partial\varphi^2} = \frac{\partial^2 E_1}{\partial\varphi^2} + \frac{\partial^2 E_2}{\partial\varphi^2} \\ &\approx \frac{\partial \mathbf{R}_1^T}{\partial\varphi}\mathbf{H}_1\frac{\partial \mathbf{R}_1}{\partial\varphi} + \frac{\partial \mathbf{R}_2^T}{\partial\varphi}\mathbf{H}_2\frac{\partial \mathbf{R}_2}{\partial\varphi} \\ &\quad - \mathbf{f}_1^T\frac{\partial^2 \mathbf{R}_1}{\partial^2\varphi} - \mathbf{f}_2^T\frac{\partial^2 \mathbf{R}_2}{\partial^2\varphi} \\ &\approx \Delta R^2(\Theta^T\mathbf{H}_1\Theta + \Theta^T\mathbf{H}_2\Theta - 2C_x) \propto \Delta R^2 4|C_x|\end{aligned} \tag{31}$$

$$\begin{aligned}\varphi_1 &= \frac{1}{2}\arctan\left(-\frac{2\,(\partial C_x/\partial\varphi)|_{\varphi=0}}{(\partial^2 C_x/\partial\varphi^2)|_{\varphi=0}}\right) \\ &\approx \frac{1}{2}\arctan\left(\frac{(\mathbf{f}_1(0) - \mathbf{f}_2(0))^T\cdot\Theta(0)}{2\Delta R|\,C_x|_{\varphi=0}|}\right).\end{aligned} \tag{32}$$

If φ_1 is smaller than 0.01 (or a different predefined value) no rotation and no SCF calculation are necessary, the dimer axis is aligned with the lowest curvature mode in the plane spanned by $\mathbf{N}$ and Θ.

The translational step of the dimer method is also slightly modified by using a different line search. A step with a predefined step length of 0.01 bohr or $\Delta x = 0.5\ (\mathbf{G}^T\mathbf{f}^\dagger)/|C_x|$ is chosen ($\mathbf{G}$ is the normalized search direction), the modified force is calculated and then a linear inter/extrapolation is done to find the root of the modified force. It is noted that C_x is not the curvature along the modified force but along the dimer axis, so that a Newton step is not chosen. To summarize, the improved dimer method proposed here requires three gradient calculations and one energy calculation, or four gradient calculations, instead of six gradient calculations as proposed by Henkelman and Jónsson[12] and at least four gradient calculations as suggested by Olsen *et al.*[29] In addition, the improved dimer method is significantly more accurate for calculating the minimum curvature mode and therefore much more robust than the original dimer methods[12,29] when energies and forces are subject to numerical noise.

The initial position for the midpoint of the dimer is taken as that of the highest-energy point determined by the interpolation algorithm. The dimer axis is initially aligned along the tangent to the approximate minimum-energy pathway at the highest-energy point. Even if this dimer axis is not aligned in the direction of the lowest curvature mode (the initial guess might have multiple unstable modes), it is most

TABLE I. List of investigated reactions and properties of initial starting configuration.

	Reaction investigated	No. of atoms	No. of negative eigenvalues in starting Hessian	Mode with maximum overlap with path	Overlap of initial path with imaginary mode at TS
1	$HCN \leftrightarrow HNC$	3	1	1	0.97
2	$HCCH \leftrightarrow CCH_2$	4	0	1	0.92
3	$H_2CO \leftrightarrow H_2+CO$	4	1	1	0.94
4	$CH_3O \leftrightarrow CH_2OH$	5	1	1	0.98
5	Cyclopropyl ring opening	8	1	2	0.62
6	Bicyclo[1.1.0] butane ring opening TS2	10	6	5	0.88
7	1,2-migration-(formyloxy)ethyl	10	1	6	0.21
8	*cis*-butadiene+ethylene $\leftrightarrow$ cyclohexene	16	7	5	0.53
9	*s*-tetrazine $\leftrightarrow 2HCN+N_2$	8	2	1	0.96
10	*trans*-butadiene $\leftrightarrow$ *cis*-butadiene	10	3	1	0.72
11[a]	$CH_3CH_3 \leftrightarrow CH_2CH_2+H_2$	8	1	1	0.94
12	$CH_3CH_2F \leftrightarrow CH_2CH_2+HF$	8	1	1	0.88
13	vinyl alcohol $\leftrightarrow$ acetaldehyde	7	1	1	0.98
14	$HCOCl \leftrightarrow HCl+CO$	4	1	1	0.85
15	$H_2PO_4^- \leftrightarrow H_2O+PO_3^-$	7	3	1	0.84
16	$CH_2CHCH_2-O-CHCH_2 \leftrightarrow CH_2CHCH_2CH_2CHO$	14	2	1	0.16
17	$SiH_3CH_2CH_3 \leftrightarrow SiH_2+CH_3CH_3$	11	2	1	0.98
18	$HNCCS \leftrightarrow HNC+CS$	5	1	1	0.74
19	$HCONH_3^+ \leftrightarrow NH_4^+ + CO$	7	1	1	0.71
20	Rotational TS in acrolein	8	3	1	0.74
21	$OCHNHOH \leftrightarrow HOCHNOH$	7	2	1	1.00
22	$H_2CNH \leftrightarrow HNC+H_2$	5	3	1	0.83
23	$H_2CNH \leftrightarrow HCNH_2$	5	2	1	0.96
24[a]	$HCNH_2 \leftrightarrow HCN+H_2$	5	1	1	0.86

[a]Growing-string method with 13 nodes.

likely aligned with the direction of the reaction coordinate. As in the P-RFO algorithm, the maximum dimer displacement for translation is set to 0.1 bohr.

RESULTS

The performance of the original dimer method developed by Henkelman and Jónsson[12] and later modified by Olsen *et al.*[29] was compared with the improved dimer method proposed here. This comparison was done for the standard suit of test reactions given by Baker and Chan[25] and presented in Table I. Table II shows the number of cycles and the number of gradient calculations necessary to converge from the starting configuration generated by the growing-string method to the transition state connecting reactant and product states. If a gradient calculation is done during the rotational step of the improved dimer method, independent of the line search algorithm used, the improved dimer method always converges to the right transition state and needs fewer cycles and gradients than the original dimer method. This speedup is especially pronounced if a SCF convergence criterion of 10^{-6} E_h is used for the DFT calculation. In this case, virtually all saddle-point searches fail using the original dimer method, whereas all but one [reaction (19) which has a very loose transition state] of the searches is successful using the improved dimer method. Preoptimizations with a reduced SCF accuracy are often helpful, since a SCF calculation with a convergence criterion of 10^{-7} E_h requires about 40% more SCF cycles to converge than a SCF calculation with a convergence criterion of 10^{-6} E_h. The significant speedup and robustness of the improved dimer method come from the improved rotational step. In the improved dimer method the minimum curvature in the plane spanned by $\mathbf{N}$ and $\mathbf{\Theta}$ is always found within the accuracy of the curvature calculation. In the original dimer method the minimum curvature is not always found. Owing to the finite difference approximation used to calculate the second derivative of the total dimer energy E and the numerical noise in the SCF calculation, the dimer axis is sometimes rotated in arbitrary directions. This results in significantly more cycles required to converge the initial estimate of the transition state to its final configuration and sometimes even to a failure to converge to the transition state. (An algorithm is considered "failed to converge" to a transition state, if it did not converge within 500 cycles.) This situation was encountered with a SCF accuracy of 10^{-7} E_h in the DFT calculation, for example, for the Diels-Alder reaction [reaction (8)] and the rotational isomerization of acrolein [reaction (20)].

It is important to note that the improved dimer method always converged to the correct transition state even if the initial dimer axis was not aligned with the eigenvector of the imaginary mode at the transition state. Table I shows that for the ring opening reactions (7) and (16), the overlap of the dimer axis $\mathbf{N}$ and the eigenvector for the imaginary mode at the transition state is 0.21 and 0.16, respectively. Nevertheless, the improved dimer method identifies the reaction coordinate relatively rapidly and converges to the saddle point. Table II shows that the performance of the improved dimer method with gradient calculations for dimer rotation is more or less independent of the line search algorithm used.

Table III illustrates the performance of the improved

TABLE II. Comparison of various dimer methods for finding saddle points in chemical reactions. A conjugate gradient algorithm was used for dimer rotation and translation.

	Improved dimer method gradient calculation for dimer rotation								Original dimer method with modifications from Olsen *et al.*[a]	
	Variable rotational angle Fixed translational step of 0.01 a.u.		Fixed rotational angle of 45° Variable translational step		Fixed rotational angle 45° Fixed translational step of 0.01 a.u.		Variable rotational angle Variable translational step			
System	No. of cycles[b]	No. of gradient calculations[b]	No. of cycles[b]	No. of gradient calculations[b]	No. of cycles[b]	No. of gradient calculations[b]	No. of cycles[b]	No. of gradient calculations[b]	No. of cycles[b]	No. of gradient calculations[b]
1	10 (8)	42 (34)	12 (7)	50 (30)	9 (7)	38 (30)	11 (7)	46 (30)	12 (9)	50 (38)
2	11 (9)	46 (38)	11 (19)	46 (78)	12 (9)	50 (38)	14 (22)	58 (90)	24 (failed)	98 (failed)
3	23 (16)	93 (66)	22 (23)	90 (94)	20 (21)	82 (86)	26 (29)	100 (117)	29 (70)	118 (282)
4	14 (10)	56 (40)	13 (13)	54 (54)	15 (13)	62 (54)	14 (10)	57 (40)	15 (failed)	62 (failed)
5	37 (66)	148 (266)	36 (44)	146 (178)	40 (31)	162 (126)	49 (42)	196 (169)	40 (failed)	162 (failed)
6	44 (42)	175 (169)	49 (40)	198 (162)	43 (44)	174 (178)	48 (40)	192 (161)	48 (failed)	194 (failed)
7	42 (101)	161 (400)	46 (64)	186 (258)	49 (42)	198 (170)	48 (57)	184 (227)	50 (failed)	202 (failed)
8	97 (96)	387 (383)	100 (87)	402 (348)	100 (94)	402 (378)	97 (89)	384 (355)	150 (failed)	602 (failed)
9	19 (17)	77 (69)	18 (17)	74 (70)	20 (17)	82 (70)	18 (17)	74 (69)	26 (failed)	106 (failed)
10	48 (126)	194 (550)	88 (178)	354 (714)	102 (90)	410 (362)	108 (170)	434 (682)	62 (failed)	250 (failed)
11[c]	25 (29)	96 (114)	24 (19)	98 (78)	25 (16)	102 (66)	25 (19)	99 (77)	25 (failed)	102 (failed)
12	25 (24)	101 (97)	26 (21)	106 (86)	26 (21)	106 (86)	24 (23)	94 (93)	26 (failed)	106 (failed)
13	14 (14)	56 (55)	15 (12)	62 (50)	14 (12)	58 (50)	14 (15)	55 (61)	13 (failed)	54 (failed)
14	15 (11)	61 (46)	11 (35)	46 (142)	14 (12)	58 (50)	15 (10)	61 (41)	15 (failed)	62 (failed)
15	37 (28)	148 (112)	39 (32)	158 (130)	38 (31)	154 (126)	38 (28)	152 (114)	42 (failed)	170 (failed)
16	32 (41)	129 (166)	32 (42)	130 (170)	32 (20)	130 (82)	32 (53)	130 (213)	38 (failed)	154 (failed)
17	42 (16)	163 (66)	30 (16)	122 (66)	43 (15)	174 (62)	38 (14)	144 (58)	50 (failed)	202 (failed)
18	20 (15)	82 (62)	31 (20)	126 (82)	24 (59)	98 (238)	22 (48)	68 (193)	28 (failed)	114 (failed)
19	38 (failed)	154 (failed)	46 (failed)	186 (failed)	34 (failed)	138 (failed)	43 (failed)	174 (failed)	42 (failed)	170 (failed)
20	39 (58)	158 (234)	124 (102)	498 (410)	51 (103)	206 (414)	114 (71)	457 (286)	failed to converge	
21	31 (17)	121 (70)	32 (19)	130 (78)	33 (19)	134 (78)	35 (19)	136 (77)	27 (failed)	110 (failed)
22	23 (24)	91 (96)	25 (22)	102 (90)	24 (25)	98 (102)	27 (26)	110 (105)	28 (failed)	114 (failed)
23	26 (8)	97 (34)	26 (7)	106 (30)	26 (8)	106 (34)	26 (8)	101 (33)	28 (failed)	114 (failed)
24[c]	20 (29)	81 (117)	23 (23)	94 (94)	24 (19)	98 (78)	24 (24)	95 (98)	29 (failed)	118 (failed)

[a]Reference 29.
[b]In parentheses: No. of cycles and gradients to reduce maximum gradient component to $<10^{-3}$ E_h/bohr. DFT calculation done with a SCF convergence criterion of 10^{-6} E_h.
[c]Growing-string method with 13 nodes.

TABLE III. Comparison of different line searches in the improved dimer method for finding saddle points if no extra gradient calculation is done during rotation. A conjugate gradient algorithm was used for dimer rotation and translation.

	Fixed rotational angle of 45° Fixed translational step of 0.01 a.u.			Variable rotational angle Fixed translational step of 0.01 a.u.			Fixed rotational angle of 45° Variable translational step			Variable rotational angle Variable translational step		
System	No. of cycles	No. of gradient calculations	Extra energy calculations	No. of cycles	No. of gradient calculations	Extra energy calculations	No. of cycles	No. of gradient calculations	Extra energy calculations	No. of cycles	No. of gradient calculations	Extra energy calculations
1	10	32	10	16	50	15	13	41	13	14	44	14
2	16	50	16	22	68	22	15	47	15	17	53	17
3	22	68	22	86	260	81	22	68	22	31	98	31
4	15	47	15	42	128	42	14	44	14	53	161	
5		Converged to wrong TS			Converged to wrong TS		31	95	31	102	308	98
6	45	137	45	72	218	69	50	152	50	69	209	68
7	48	146	48	232	698	217	45	137	45	157	473	150
8	100	302	100	352	1058	334	99	299	99		Failed to converge	
9	19	59	19	21	65	20	18	56	18	20	62	20
10	272	818	272		Failed to converge		81	245	81	46	140	46
11[a]	25	77	25	66	200	64	26	80	26	50	152	46
12	24	74	24	74	224	73	26	80	26	82	248	78
13	14	44	14	32	98	32	15	47	15	37	113	37
14	14	44	14	13	41	12	13	41	13	19	59	18
15	37	113	37	83	251	77	37	113	37	75	227	69
16	29	89	29	198	596	192	31	95	31	45	137	44
17	42	128	42	95	287	88	42	128	42	123	371	119
18	20	62	20	29	89	29	25	77	25	21	65	20
19	37	113	37	48	146	47	40	122	40	44	134	42
20	109	329	109	52	158	52	50	152	50	89	269	89
21	32	98	32	65	197	60	29	89	29	62	188	60
22	28	86	28	96	290	90	27	83	27	38	116	38
23	33	101	33	138	416	129	26	80	26	58	176	55
24[a]	25	77	25	119	359	110	22	68	22	36	110	34

[a]Growing-string method with 13 nodes.

TABLE IV. Comparison of improved dimer methods with steepest descent or with conjugate gradient algorithm for dimer rotation and translation. A gradient calculation was done during rotation. The improved dimer method was used with a variable initial rotation angle and a fixed initial translational step of 0.01 bohr.

	CG for rotation SD for translation		SD for rotation SD for translation		SD for rotation CG for translation		CG for rotation CG for translation	
System	No. of cycles	No. of gradient calculations	No. of cycles	No. of gradient calculations	No. of cycles	No. of gradient calculations	No. of cycles	No. of gradient calculations
1	24	97	27	110	8	34	10	42
2	25	102	17	70	12	50	11	46
3	27	103	30	122	20	81	23	93
4	37	129	33	127	12	49	14	56
5	50	191	Converged to wrong TS		28	114	37	148
6	85	336	85	342	43	174	44	175
7	110	404	115	456	42	170	42	161
8	202	781	206	826	107	430	97	387
9	48	189	44	178	16	66	19	77
10	145	581	149	598	68	274	48	194
11[a]	52	183	45	179	24	97	25	96
12	57	222	49	198	28	113	25	101
13	31	117	26	105	18	73	14	56
14	32	130	32	130	17	70	15	61
15	69	274	121	485	41	166	37	148
16	72	290	77	310	29	118	32	129
17	58	226	62	248	38	154	42	163
18	37	150	39	158	17	70	20	82
19	134	536	124	498	32	130	38	154
20	123	494	84	338	48	194	39	158
21	58	228	62	250	30	122	31	121
22	50	200	50	202	24	98	23	91
23	79	316	111	446	27	109	26	97
24[a]	43	165	44	178	29	118	20	81

[a]Growing-string method with 13 nodes.

dimer method if an energy calculation is done to determine the Fourier coefficients in Eq. (24). Using an energy calculation instead of a gradient calculation saves one gradient calculation per cycle but also reduces the accuracy of the curvature calculation. The performance of the algorithm depends on the method used to determine the Fourier coefficients in Eq. (24). If a large enough angular rotation, e.g., 45°, is done, the Fourier coefficients in Eq. (24) are found with high accuracy. On the other hand, if a variable rotation of the dimer axis is done, Eq. (32), a small rotation angle is sometimes chosen, resulting in a curvature calculation that is not sufficiently accurate to determine the Fourier coefficients and the curvature minimum. The dimer separation ΔR then has to be reduced, but this is not possible because it increases the impact of the numerical noise in the SCF calculation. As a result, the improved dimer method with variable rotation angle rotates the dimer axis in a manner similar to that used in the original dimer method proposed by Henkelman and Jónsson,[12] which can sometimes result in the selection of arbitrary directions and the requirement of more than twice as many cycles to converge to the transition state. The improved dimer method without gradient calculation during rotation, but using a fixed rotation angle of 45° in the rotational step, performs similar to the improved dimer method with gradient calculation. For a level of theory where the gradient of the potential energy can be obtained very rapidly with a well-converged wave function using the generalized Hellmann-Feynman theorem, this modification to the improved dimer method is not recommended. The reduction in the number of gradient calculation per iteration cycle is outweighed by the reduced accuracy in curvature calculation and the corresponding risk of not finding the minimum curvature mode in the plane spanned by $\mathbf{N}$ and Θ. On the other hand, if the gradient calculation has to be performed numerically, this modification to the improved dimer method is recommended.

Henkelman and Jónsson[12] suggest the use of a conjugate gradient (CG) algorithm for the translational and rotational steps in the original dimer method. Table IV illustrates the benefit of using such an algorithm. The use of the CG algorithm of Polak-Ribière[32] for the translational step greatly improves the dimer method by more than halving the number of cycles and gradient calculations needed to converge to a saddle point. The conjugate gradient algorithm for the rotational step[12] on the other hand does not significantly improve the performance of the dimer method over a steepest descent (SD) step. The reason for the lack of improvement in the CG algorithm originates from not generating a set of conjugate directions. After every rotation a translational step is done such that the coordinates $\mathbf{R}_0$, $\mathbf{R}_1$, and $\mathbf{R}_2$ change and no set of conjugate directions for rotation can be generated. This

TABLE V. Comparison of improved dimer method with P-RFO algorithm for finding saddle points in chemical reactions.

	P-RFO method				Improved dimer method[a]			
	Hessian matrix updating formula from Bofill follow lowest mode				Gradient calculated for rotation Variable rotational angle		No gradient calculated for rotation Fixed rotational angle of 45°	
	Analytical Hessian		Numerical Hessian		Fixed translational step of 0.01 a.u.		Variable translational step	
System	No. of cycles[b]	Time (s)	No. of cycles[b]	Time (s)	No. of cycles[b]	Time (s)	No. of cycles[b]	Time (s)
1	5 (4)	35	5 (4)	72	10 (7)	81	13 (10)	100
2	7 (6)	66	7 (6)	118	11 (9)	126	15 (12)	160
3	13 (12)	87	13 (12)	141	23 (19)	270	22 (18)	249
4	7 (6)	113	7 (6)	191	14 (11)	250	14 (11)	254
5	Converged to wrong TS[c]		Converged to wrong TS[c]		37 (28)	2 802	31 (27)	2 319
6	89 (81)	4185	96 (92)	5449	44 (41)	5 587	50 (42)	5 904
7	30 (27)	3336	28 (23)	4864	42 (30)	8 148	45 (27)	9 611
8	Failed to converge[c]		Failed to converge[c]		97 (89)	42 201	99 (91)	43 263
9	12 (9)	1594	12 (10)	2445	19 (17)	3 268	18 (15)	3 236
10	27 (19)	1777	21 (15)	2219	48 (34)	5 543	81 (48)	9 061
11[d]	16 (14)	428	16 (14)	658	25 (18)	922	26 (18)	992
12	19 (15)	691	20 (15)	1001	25 (21)	1 510	26 (22)	1 499
13	7 (6)	353	7 (6)	579	14 (11)	681	15 (11)	725
14	8 (7)	128	8 (7)	212	15 (9)	338	13 (9)	294
15	29 (26)	1392	29 (24)	1806	37 (31)	3 560	37 (34)	3 517
16	15 (11)	4862	11 (10)	7681	32 (21)	11 445	31 (20)	11 308
17	13 (10)	1155	15 (9)	1929	42 (14)	3 712	42 (14)	3 862
18	8 (6)	268	7 (6)	399	20 (16)	902	25 (18)	1 104
19	39 (30)	653	23 (21)	637	38 (26)	1 380	40 (28)	1 438
20	28 (26)	1202	Failed to converge		39 (23)	3 310	50 (26)	4 095
21	10 (9)	552	11 (9)	877	31 (17)	1 902	29 (18)	1 835
22	24 (22)	183	23 (22)	256	23 (21)	371	27 (23)	418
23	11 (8)	128	11 (8)	206	26 (8)	390	26 (8)	401
24[d]	20 (17)	171	20 (17)	248	20 (19)	343	22 (18)	367

[a]CG for translation, CG for rotation.
[b]In parentheses: number of cycles to reduce maximum gradient component to $<10^{-3}$ E_h/bohr.
[c]Reactions (5) and (8) failed to converge to right TS even if Hessian is calculated in every cycle.
[d]Growing-string method with 13 nodes.

problem does not occur for the translational step, since the dimer midpoint is not moved during the rotational step.

Table V compares the number of cycles and the time (for a single processor on an Opteron 848, 2.2 GHz computer) necessary to converge to a saddle point connecting reactant and product states using the improved dimer method and the P-RFO method. A saddle point is found with the P-RFO algorithm if only one Hessian calculation (analytical or numerical) is done for the starting configuration and the Hessian updating scheme of Bofill[11] is used thereafter. A regular calculation of the second derivative matrix is not required. If the P-RFO method converges to a transition state it needs on average 25%–30% fewer cycles and about 35% (numerical Hessian calculation) to 55% (analytical Hessian calculation) less time than the improved dimer method. The dimer method is not quadratic convergent close to the saddle point as is the P-RFO method, which uses the second derivative matrix of the potential energy. If the convergence criterion to find a transition state is relaxed to a maximum gradient component of 10^{-3} E_h/bohr, the improved dimer method requires 10%–20% more cycles to converge to the transition state. The P-RFO method is superior to the improved dimer method for problems in which the starting Hessian matrix has one and only one negative eigenvalue and this eigenvector for this mode is aligned with the reaction coordinate (see Table I). However, the P-RFO method has problems finding the correct transition state if an eigenvector not corresponding to the lowest eigenvalue is aligned with the reaction coordinate. Table VI illustrates that this problem does not fully disappear if a mode-following-type algorithm is used and the Hessian matrix is calculated regularly. The P-RFO method failed to converge to the right transition state for reactions (5) and (8). The improved dimer method, on the other hand, has fewer problems with multiple negative eigenvalues and a reaction coordinate not corresponding to the eigenvector with the lowest eigenvalue. As long as the initial dimer orientation has significant overlap with the reaction coordinate the improved dimer method has a larger convergence radius than the P-RFO method and is therefore more robust.

In saddle-point searches involving heavy atoms, in particular, transition metals, the potential-energy surface is often shallow and the second derivative matrix has to be calculated regularly in the P-RFO algorithm. Under these circumstances, it is possible to determine approximately the system size above which the improved dimer method outperforms the P-RFO method for problems where both algorithms converge to the correct transition state. Assuming the Hessian matrix is calculated every X cycles, a Hessian calculation

TABLE VI. Performance of a mode-following P-RFO algorithm where the mode-followed uphill is the one that has the greatest overlap with the mode-followed uphill in the previous step. Data just illustrated if the mode-followed uphill is at least one time not the one with the lowest eigenvalue.

	Analytical Hessian				Numerical Hessian			
	1 Hessian calculated		Every fifth cycle Hessian calculated		1 Hessian calculated		Every fifth cycle Hessian calculated	
System	No. of cycles	No. of gradient calculations	No. of cycles	No. of gradient calculations	No. of Hessian calculations	No. of cycles	No. of gradient calculations	No. of Hessian calculations
5	Failed to converge		Failed to converge			Converged to wrong TS	74	15
6	114	114	Converged to wrong TS			Failed to converge	Failed to converge	
7	Failed to converge		Converged to wrong TS			Failed to converge	Converged to wrong TS	
8	Failed to converge		Failed to converge			Failed to converge	Failed to converge	

takes about as much time as $3n$ gradient calculations (C_1 symmetry), where n is the number of atoms in the system, and the improved dimer method needs $\alpha\%$ (about 35%–45% using tight convergence criteria and about 10%–20% for loose convergence criteria; in fact, α is also a function of X) more cycles to converge, the improved dimer method is superior if

$$X + 3n > 4X\left(1 + \frac{\alpha}{100}\right)$$

$$n > X\left(1 + \frac{4}{3}\frac{\alpha}{100}\right)$$

$$n > 1.5X \text{ (tight convergence criteria)}$$

$$n > 1.2X \text{ (loose convergence criteria)}. \quad (33)$$

For example, if the Hessian matrix is calculated every 10th cycle the improved dimer method should be used for systems with more than 15 atoms (tight convergence criteria) or 12 atoms (loose convergence criteria). It is also noted that the improved dimer method has a larger convergence radius if a reasonable initial path tangent guess is available. This is especially relevant for problems with many negative eigenvalues in the Hessian matrix at the initial configuration.

For all of the test calculations reported here, the algorithm spends more than 99.9% of the time in the QM program package. The determination of the eigenvalues and eigenvectors of the Hessian matrix in the P-RFO algorithm is never rate limiting. For very large systems with an analytical (cheap) potential this step can become rate limiting and the dimer methods will always outperform algorithms based on a second derivative matrix.[29]

CONCLUSIONS

An improved dimer method for finding transition states is presented and compared with the original dimer method of Henkelman and Jónsson[12] as modified by Olsen *et al.*[29] and with the well-established partitioned rational function optimization method.[8] In all cases, an interpolation method is used to obtain a reasonable guess for the transition state and the path tangent at the starting configuration. The improved dimer method with gradient calculation during rotation requires significantly fewer cycles and is much more robust than the original dimer method when energies and forces are subject to numerical noise. If a gradient calculation takes significantly longer than an energy calculation, e.g., the gradient has to be calculated numerically, the gradient calculation during the rotational step can be replaced with an energy calculation. A conjugate gradient algorithm for the translational step in the dimer method improves its performance; however, a conjugate gradient approach for the rotational step does not significantly improve the algorithm over a steepest descent direction for rotation. After every rotation a translation is done and no set of conjugate directions for rotation is generated. Depending how tight the convergence criterion is chosen the improved dimer method requires about 1.10–1.45 more cycles to converge to a saddle point

than a P-RFO algorithm using a Hessian updating scheme. As a result of the better scaling with system size, the improved dimer method is recommended relative to the P-RFO method for systems with more than $1.5\times X$ atoms ($1.2\times X$ atoms for a loose convergence criterion), where the Hessian matrix is calculated in the P-RFO method every X cycle. A mode-following algorithm based on the P-RFO can fail to find the right transition state if the reaction coordinate does not correspond to the eigenvector with the lowest eigenvalue. Under similar circumstances, the improved dimer method converges to the correct transition state, provided a good initial guess for the path tangent is available.

The FORTRAN code for the improved dimer method is available and can be supplied on request.

ACKNOWLEDGMENTS

The authors appreciate helpful discussions with Dr. Baron Peters and Dr. Graeme Henkelman. This work was supported by the Methane Conversion Cooperative funded by BP, the "Fonds der chemischen Industrie," and the Max-Buchner-Forschungsstiftung.

[1] C. Wert and C. Zener, Phys. Rev. **76**, 1169 (1949).
[2] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).
[3] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).
[4] J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, J. Phys. Chem. **87**, 2745 (1983).
[5] A. Banerjee, N. Adams, J. Simons, and R. Shepard, J. Phys. Chem. **89**, 52 (1985).
[6] J. Nichols, H. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. **92**, 340 (1990).
[7] D. J. Wales, J. Chem. Phys. **91**, 7002 (1989).
[8] J. Baker, J. Comput. Chem. **7**, 385 (1986).
[9] B. A. Murtagh and R. W. H. Sargent, Comput. J. **13**, 185 (1970).
[10] M. J. D. Powell, *Nonlinear Programming* (Academic, New York, 1970).
[11] J. M. Bofill, J. Comput. Chem. **15**, 1 (1994).
[12] G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999).
[13] R. Malek and N. Mousseau, Phys. Rev. E **62**, 7723 (2000).
[14] G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).
[15] G. Mills and H. Jónsson, Phys. Rev. Lett. **72**, 1124 (1994).
[16] B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, J. Chem. Phys. **120**, 7877 (2004).
[17] W. Ren, Ph.D. thesis, Department of Mathematics, New York University, 2002.
[18] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[19] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).
[20] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[21] O. Treutler and R. Ahlrichs, J. Chem. Phys. **102**, 346 (1995).
[22] A. Schäfer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).
[23] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett. **162**, 165 (1989).
[24] R. Ahlrichs and M. v. Armin, in *Methods and Techniques in Computational Chemistry: METECC-95*, edited by E. Clementi and G. Corongiu (STEF, Calgiari, 1995).
[25] J. Baker and F. Chan, J. Comput. Chem. **17**, 888 (1996).
[26] M. Page and J. W. McIver, Jr., J. Chem. Phys. **88**, 922 (1988).
[27] J.-Q. Sun and J. Ruedenberg, J. Chem. Phys. **99**, 5269 (1993).
[28] C. Gonzalez and H. B. Schlegel, J. Chem. Phys. **95**, 5853 (1991).
[29] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, J. Chem. Phys. **121**, 9776 (2004).
[30] A. Heyden and F. J. Keil, AIChE Annual Meeting, paper 566h (San Francisco, CA, 2003).
[31] A. Heyden, B. Peters, A. T. Bell, and F. J. Keil, J. Phys. Chem. B **109**, 1857 (2005).
[32] E. Polak and G. Ribière, Rev. Fr. Inform. Rech. Oper. **16**, 35 (1969).