JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Improving Upon String Methods for Transition State Discovery

Hugh Chaffey-Millar,[†] Astrid Nikodem,[‡] Alexei V. Matveev,[‡] Sven Krüger,[‡] and Notker Rösch*[,‡]

[†]Molekulare Katalyse, Department Chemie, Technische Universität München, 85747 Garching, Germany

[‡]Department Chemie & Catalysis Research Center, Technische Universität München, 85747 Garching, Germany

Supporting Information

ABSTRACT: Transition state discovery via application of string methods has been researched on two fronts. The first front involves development of a new string method, named the Searching String method, while the second one aims at estimating transition states from a discretized reaction path. The Searching String method has been benchmarked against a number of previously existing string methods and the Nudged Elastic Band method. The developed methods have led to a reduction in the number of gradient calls required to optimize a transition state, as compared to existing methods. The Searching String method reported here places new beads on a reaction pathway at the midpoint between existing beads, such that the resolution of the path discretization in the region containing the transition state grows exponentially with the number of beads. This approach leads to favorable convergence behavior and generates more accurate estimates of transition states from which convergence to the final transition states occurs more readily. Several techniques for generating improved estimates of transition states from a converged string or nudged elastic band have been developed and benchmarked on 13 chemical test cases. Optimization approaches for string methods, and pitfalls therein, are discussed.

1. INTRODUCTION

One of the most important tasks in computational chemistry is the discovery of geometries and energies of transition states (TSs). As is common practice, with the term TS we refer to a saddle point on the potential energy surface of a chemical system. In general, one searches for the highest point on a minimum energy pathway between the reactant and product minima of a chemical reaction.[1,2]

The typical procedure for locating a TS is as follows.[1] Starting from an initial estimate of the TS's geometry, a gradient driven optimization is performed to locate the nearest first-order saddle point, that is, a stationary point on the potential energy surface (PES) that has negative curvature in exactly one direction and positive curvature in all others. Mathematically speaking, if $\mathbf{x}_{TS}$ is the vector of internal or Cartesian coordinates of the TS, and $E(\mathbf{x})$ is the energy of a molecule with coordinates $\mathbf{x}$, then $\mathbf{x}_{TS}$ satisfies $dE/dx_i = 0$ and the matrix $d^2E/dx_i dx_j$ has exactly one negative eigenvalue.

The initial estimate of a TS might be formed using, for example, chemical intuition or some automated procedure. Many automated procedures invoke knowledge of the geometries of the reagents and search first for some approximation of the entire reaction path, from which an approximation of the TS is formed.[1,2] These automated procedures are addressed in the current contribution, together with a refinement of the initial TS estimate using a local optimization method that has previously been reported.[3−7]

One category of automated TS discovery techniques is known as the chain-of-state methods.[1] In these methods, the reaction pathway is approximated via discretization; that is, a chain of N geometries (or "beads") is formed by interpolating between the reactant and product geometries. The most basic of these methods, known as the Linear Synchronous Transit (LST) method,[1] involves forming a series of equally spaced beads with geometries interpolated linearly. The bead with the highest energy is used as the initial estimate for subsequent refinement via, for example, a Newton method or the dimer method.[3]

However, such a procedure for selecting the TS estimate often is too simple. An improved approach results if one optimizes the beads in all directions that are perpendicular to the reaction path. Specifically, if the tangent to the reaction path at bead i can be assumed to be approximately $\tau_i = 1/2[(\mathbf{x}_i - \mathbf{x}_{i-1}) + (\mathbf{x}_{i+1} - \mathbf{x}_i)]$, then one optimizes bead i in all directions perpendicular to τ_i. It is desirable that the beads sample the reaction path in the most relevant manner possible, by suitably constraining the spacing of the beads along the reaction pathway. Therefore, during their optimization, the beads should not be allowed to accumulate at the minima or lose the desired spacing constraint.

A distinguishing feature of all chain-of-state methods is the mechanism by which this spacing is achieved. The Nudged Elastic Band (NEB) method[8−10] augments the force acting on each bead with spring-like forces that act only along the path. The string method and derivatives thereof[11−13] use a spline fitted through the beads to facilitate measurement of the distance between them as well as repositioning to satisfy the spacing constraint.

There is ongoing research into methods for locating TSs and/or reaction trajectories. The original string method[11] continues to be discussed with various improvements,[2,4,12−19] as does the NEB method.[4,9,10,20−23] Types of chain-of-state methods that are neither string nor NEB variants have been explored,[24,25] and finally, methods which do not rely on a discretization of the reaction pathway, such as the Fast Marching method,[15,26−28] continue to receive attention.

In the current contribution, the NEB and various string methods are applied to the task of automated TS searching. Note that this is a distinct task from the discovery of entire reaction

Received: September 13, 2011
Published: January 4, 2012

   dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777−786

pathways, since only the highest region of the pathway is of interest. Central objectives of all research into chain-of-state methods are (a) the reduction of the computational effort required to converge the band or string, while at the same time, (b) improving the accuracy of the TS estimate, since a more accurate TS estimate will save time and increase the likelihood of success of the subsequent local TS search. Since the computer time required is directly related to the number of times the energy, $E(\mathbf{x})$, and energy gradients, $dE/d\mathbf{x}$, are calculated, the current study focuses on two fronts: (1) methods of reducing the number of beads, involving the development of a new type of string method, and (2) procedures for generating improved estimates of the TS geometry based upon the string/band.

The new string method is similar to the previously described Growing String method,[13] where the string is grown "inwards" from the ends of the path, but differs in the manner in which beads are placed during growth. In fact, each new bead is added to the midpoint of the bead interval most likely to contain the TS. The pathway interpolation methods examined use information from the beads in the vicinity of the TS with the objective of generating a TS estimate that converges more readily to the correct TS than would the highest bead.

The structure of the paper is thus as follows: (1) the NEB and string methods are briefly summarized, and a new string method is proposed, including development of an optimizer for string methods; (2) the interpolation methods are described; (3) the developed methods are benchmarked using a library of 13 molecular potential energy surfaces derived from seven chemical reactions, and (4) the results are summarized.

2. DESCRIPTION OF METHODS

2.1. NEB and String Methods. We assume that, for any position on a molecule's PES with coordinates $\mathbf{x} = (x_1,...,x_D)^T$, some code provides the energy of the molecule, $E(\mathbf{x})$, and the corresponding nuclear displacement gradient, $\mathbf{g}(\mathbf{x}) = (\partial E/\partial x_1,...,\partial E/\partial x_D)^T$, at $\mathbf{x}$, where D is the number of coordinates describing the molecule.

In the optimization of a NEB, the effective force (i.e., the negative of the gradient) acting on each bead i comprises two contributions: one originating from the component of the energy gradient perpendicular to the reaction path, $\mathbf{g}_i^{\perp}$, that corresponds to a projection of chemical forces, and one parallel to the path, $\mathbf{g}_i^{\parallel}$, that arises from artificial elastic forces designed to keep the beads evenly spaced. These contributions are described in the Supporting Information and also elsewhere.[4,9,10,20−23] Whether each bead separately or the entire set of beads system is treated by a single optimizer instance varies between implementations.[23] In the current work, optimization progress is defined for the vector $\mathbf{X}$ composed of the coordinates of all N beads $\mathbf{X} = (\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)^T$ using the overall effective energy gradient, $\mathbf{G}(\mathbf{X}) = (\mathbf{g}_1^{\mathrm{NEB}},...,\mathbf{g}_N^{\mathrm{NEB}})^T$.

The key aspect of the string method which differentiates it from the NEB method is the fact that splines are utilized for calculating the tangent of the path and for enforcing a desired bead spacing. We briefly describe how this is realized in the current work.

A spline $\phi_j(t)$ is fitted along each molecular coordinate j, giving an overall pathway parametrization $\Phi(t) = (\phi_1(t), \phi_2(t),...,\phi_D(t))$, where $t \in [0,1]$. The vectors $\Phi(0)$ and $\Phi(1)$ are the coordinates of the reactant and products, respectively, and the discretized points $\mathbf{x}_i = \Phi(t_i)$ along the reaction pathway occur at t_i such that $t_{i+1} > t_i$. The choice of the values of t_i is discussed later.

This definition of a spline through the points facilitates calculation of the path tangent, $\boldsymbol{\tau}$, and provides a means to maintain a predetermined spacing of the beads via repositioning of the beads along the spline path. A string length parameter, $s(t)$, is now defined as a solution of the differential equation $ds = |d\Phi/dt|dt$ at any point of the string. The unit tangent vector is then defined as $\tau(s) = d\Phi/ds$. Similar to the NEB method, the component of the gradient that is perpendicular to the path at each bead i is evaluated via projection:

$$\mathbf{g}_i^{\perp} = \mathbf{g}(\mathbf{x}_i) - \tau_i(\tau_i \cdot \mathbf{g}(\mathbf{x}_i)) = (I - \tau_i\tau_i^T)\mathbf{g}(\mathbf{x}_i) \quad (1)$$

At odds with the NEB approach, the desired spacing of the beads along the string is achieved by directly adjusting bead separations as measured by $s(t)$, as discussed in the Supporting Information.

2.2. The New "Searching String" Method. The Growing String methods represent an approach to reduce the number of gradient calls required to converge the pathway.[13,18] In this method, not all beads are present for the entire optimization procedure but instead are added during optimization in such a manner that the string "grows" in from its ends (i.e., the reactants and products) toward the TS. In related strategies, termed hybrid and substring methods, an initial string that is converged using a lower-level of quantum chemical theory leads to a second string that only spans the TS containing region of the reaction pathway. For constructing this second string, a higher level of quantum chemical theory may be employed, or a greater bead density and/or tighter convergence criteria.[16,17,29] These and other authors[14] used PES interpolation methods[30,31] to skip expensive electronic structure calculations for geometries that are sufficiently close to previously calculated ones. Perceived advantages in having a string grow in from its ends are (a) that computer time is only spent on midstring beads once the string is relatively close to the reaction pathway and the midstring beads are thus more likely to represent physically realistic molecular geometries and (b) that a reaction pathway can be discovered even when it differs greatly from linearity.[13]

In the current contribution, we propose a new string growth method, the Searching String method. It proceeds as follows:

1. An initial string of four beads is converged.
2. A single additional bead is placed at the midpoint between the two beads most likely to contain the TS (discussed later).
3. The new string is converged.
4. If the maximum number of beads has been reached, finish; otherwise, repeat from step 2.

This method guarantees that the greatest bead density is close to the TS and also reduces the amount of computational effort spent converging the whole reaction pathway. It bears resemblance to the weighted arc method.[32]

To decide which two beads are likely to bracket the TS, the projection of the gradient $\mathbf{g}(\mathbf{x}) = dE/d\mathbf{x}$ onto the tangent $\tau = d\mathbf{x}/ds$ is calculated for every bead, delivering the gradient along the reaction pathway, dE/ds. If the energies and dE/ds values of two adjacent beads, i and $i + 1$, satisfy

$$E(\mathbf{x}_i) > E(\mathbf{x}_{i+1}) \text{ and } dE/ds|_i > 0$$
$$\text{or}$$
$$E(\mathbf{x}_i) < E(\mathbf{x}_{i+1}) \text{ and } dE/ds|_{i+1} < 0 \quad (2)$$

then they are deemed to be candidates for bracketing a TS estimate as well as the next bead to be created.

dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777−786

2.3. Reaction Path Interpolation. In many previous studies,[13,16,18,23] the highest bead from a converged band was used as the starting point for a subsequent local TS search. However, it has been shown[9,17] that interpolated positions along a string or band that do not correspond directly to beads can make more appropriate starting points.

There also exist methods to move the highest bead uphill toward the TS.[9,17,32] For example, the climbing image method,[9] relevant to NEBs, applies additional spring forces to the highest bead to effect this. It is assumed that this highest energy bead has a greater chance of being close to the actual TS. Some authors found that, in contrast to methods without a climbing image, difficulties were experienced when applied to rough PESs.[32]

Henkelman and Jónsson have used cubic polynomials fitted to converged NEBs to obtain a more accurate estimate of the reaction pathway and hence TS.[10] Interpolation methods have also been used to extract the best possible TS estimate from a string.[17] A spline, in addition to being used to parametrize the strings' beads, was used to build a reaction energy profile, $E_{spline}(s)$, by fitting to the beads' energies (see for example the dashed curve in Figure 1).[17] The pair of beads, i and $i + 1$, which

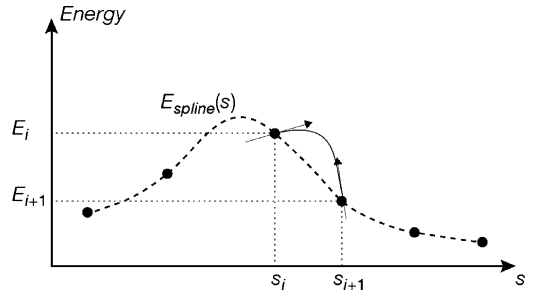


Figure 1. Schematic representation of various methods of interpolating between points on a discretized reaction pathway parametrized by the reaction coordinate s to generate a TS estimate. Except for the *Highest Bead* method, a cubic spline is commonly fitted through the beads' molecular degrees of freedom to interpolate over molecular coordinates. The *Spline* and *Weighted Average* methods use an additional spline, $E_{spline}(s)$ (dashed line), fitted through the beads' energies to find the value of s that maximizes $E_{spline}(s)$, and thus the beads which may bracket the TS. The *Spline and Polynomial* method uses the slope dE/ds at each bead, to determine which beads may bracket the TS, and a cubic polynomial, $E_{poly}(s)$ (solid line), fitted to the two energy values and slopes, dE/ds, of the beads s_i and s_{i+1}.

satisfies $s_i < s_{max} < s_{i+1}$ for some s_{max} that maximizes $E_{spline}(s)$, was deemed to bracket an improved TS estimate as compared to choosing the highest bead. On the basis of this assumption, the TS estimate was calculated as a weighted average of the geometries of beads i and $i + 1$, i.e., $\mathbf{x}_{TS} \approx w\mathbf{x}_i + (1 - w)\mathbf{x}_{i+1}$, where w is related to the corresponding normalized arc length.[17] We shall refer to this method as the *Weighted Average* method.

A possible problem with the *Weighted Average* method might be that the pair of beads bracketing the maximum of the fitted energy profile, $E_{spline}(s)$, does not necessarily bracket the best possible TS estimate. Additionally, because the weight w comes from the location of the spline maximum, its value will depend on the energy of other points than those bracketing the TS, i and $i + 1$, whereas it is the curvature of the PES at beads i and $i + 1$ that is most directly relevant to the TS. This would especially be a problem for low bead densities.

In addition to the *Weighted Average* method, we have implemented several others for comparison. In a method to which we will refer as the *Pair Average* method, the calculated dE/ds is used to determine which beads bracket the TS; however, the TS estimate is formed as the average of the geometries of the bracketing beads: $\mathbf{x}_{TS} \approx (\mathbf{x}_i + \mathbf{x}_{i+1})/2$. In the so-called *Spline* method, as in the *Weighted Average* method, a spline, $E_{spline}(s)$, is used to represent the energy profile, and the value $s = s_{max}$ which maximizes $E_{spline}(s)$ is used to generate the TS estimate as $\mathbf{x}_{TS} \approx \Phi(s_{max})$. See the dashed curve in Figure 1.

In the present work, a new interpolation method was explored, wherein gradient information from the PES is used: first, to find the beads that are likely to bracket the best TS estimate and, second, to generate a TS estimate from these beads. This is similar to the method described in the Appendix of ref 10, the essential difference being that that method was applied in a manner quite specific to the NEB approach.

Recall that for the converged string, the gradient of the energy, $dE/d\mathbf{x}$ is known for every bead. The same criteria that were earlier used to decide whether two beads bracket the TS (eq 2) can be used again. Further, for two such beads i and $i + 1$, it is possible to fit a cubic polynomial, $E_{poly}(s)$, through the points $(E(\mathbf{x}_i),s_i)$ and $(E(\mathbf{x}_{i+1}),s_{i+1})$ as well as the derivatives $dE/ds|_i$ and $dE/ds|_{i+1}$. The maximum of this cubic curve at a value $s = s_{max}$ yields an approximate location of the TS, $\mathbf{x}_{TS} \approx \Phi(s_{max})$, based on the spline vector Φ and the information on the local curvature and energy. We will refer to this approach as the *Spline and Polynomial* method. It is depicted schematically as the solid curve in Figure 1.

For methods that interpolate the reaction pathway in combination with the NEB method, a spline was fitted to the beads' degrees of freedom in the same manner as described for the string methods.

Thus, the following five methods for generating a TS estimate were benchmarked (see Table 1): (1) *Highest Bead* method, (2) *Spline* method, (3) *Weighted Average* method, (4) *Pair Average* method, and (5) the *Spline and Polynomial* method. As a pictorial summary of both the new Searching String method and the *Spline and Polynomial* interpolation method, Figure 2 depicts the discretized reaction pathways from 16 iterations of the Searching String method as it is grown from four to six beads.

Table 1. Summary of Methods for Generating a Transition State (TS) Estimate from a Converged String or (Nudged Elastic) Band

method	transition state (TS) estimate ($\approx \mathbf{x}_{TS}$)[a]
Highest Bead	$\mathbf{x}_{TS} \approx \mathbf{x}_i\|_{Emax}$
Spline	$\mathbf{x}_{TS} \approx \Phi(s_{max})$, where s_{max} maximizes $E_{spline}(s)$, cubic spline representation of the reaction energy profile
Weighted Average	$\mathbf{x}_{TS} \approx w\mathbf{x}_i + (1 - w)\mathbf{x}_{i+1}$, a weighted average of the coordinates of the beads bracketing the TS[b]
Pair Average	$\mathbf{x}_{TS} \approx (\mathbf{x}_i + \mathbf{x}_{i+1})/2$, a (nonweighted) average of the coordinates of the TS bracketing beads
Spline and Polynomial	$\mathbf{x}_{TS} \approx \Phi(s_{max})$, where s_{max} maximizes $E_{poly}(s)$, a cubic polynomial representation of the reaction energy profile for the two beads deemed to bracket the transition state

[a]See text for definitions of nomenclature. [b]Ref 17.

3. OPTIMIZER

3.1. Review of Optimization Approaches. A crucial aspect of the implementation of any chain-of-state method is the choice of the algorithm for optimization of the beads. The optimization used with the string methods is a special case because there are no spring forces to maintain the desired bead

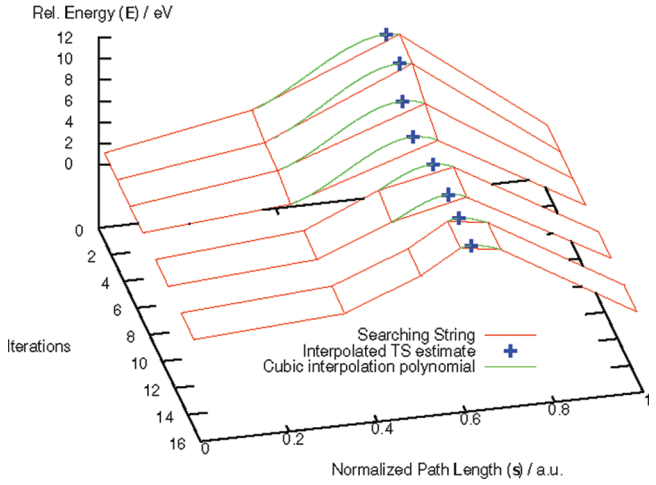


Figure 2. Evolution of a searching string optimization of a Diels–Alder reaction using **Z**-matrix coordinates (reaction III in Table 2). Starting with four beads, the string is grown until it contains six beads. The green curves are cubic polynomials, $E_{poly}(s)$, fitted to the energies of two beads and the gradient, dE/ds, at this position. This constitutes the *Spline and Polynomial* interpolation method (see text), and the maximum of the resulting cubic polynomial, indicated with the blue cross, corresponds to the value of the reaction coordinate, s, that is used to generate a transition state estimate. Only every second iteration is depicted.

spacing. A number of optimization approaches have been explored for NEB calculations,[33] including the steepest descent, conjugate gradient,[5] Fast Inertial Relaxation Engine (FIRE),[34] and quasi-Newton method with a Broyden Fletcher Goldfarb Shanno (BFGS) Hessian update.[5] The best results were reported when a single instance of the limited memory variant[35] of a BFGS (L-BFGS) optimizer was applied globally to the entire NEB, rather than having a separate instance for each bead. In the present work, we employed this approach for NEB optimizations, the L-BFGS algorithm being the same as that reported previously[35,36] with the exception of special means when the energy increases as a result of an optimization step. In this case, instead of taking a quasi-Newton step, the optimizer is instructed to step to the minimum along the line between the lowest two energies found so far. If no such minimum exists, a quasi-Newton step is tried but with a halved maximum step size. If, after three such "backtracks", the lowest energy found so far has not been improved upon, the Hessian is reset. This approach has been found to prevent divergence during optimization of some test cases.

For string methods, a variety of approaches have been tried. For example, the original Growing String method[13] used a line search along the steepest descent direction to determine a step length in that direction. Other authors applied the conjugate gradient method[16,17,29] in conjunction with a Shepard PES interpolation methodology.[15,29,30] This avoids expensive electronic structure calculations and instead uses previously calculated points to generate interpolated $\mathbf{g}(\mathbf{x})$ and $E(\mathbf{x})$ values, provided $\mathbf{x}$ was close enough to previously computed values. Recently, a special purpose optimizer was described, forming the so-called Quadratic String (QS) method.[14] This will be discussed later in relation to the development of the optimizer used in the current work. A quasi-Newton (BFGS) optimizer has also been used for the String method;[23] however, as described below, this approach is unlikely to be optimal.

In the present work, a novel optimization strategy was developed to address certain shortcomings when a quasi-Newton optimizer is directly applied to a string. This new method represents a simplification of the QS method,[14] which has certain features not strictly required for the specific task of transition state discovery. First, the problems associated with the quasi-Newton approach are explained, followed by the details of the new approach.

Recall that in quasi-Newton methods, the function $E(\mathbf{x})$ to be minimized is approximated using a second-order Taylor approximation in combination with an approximate Hessian, $\mathbf{H}$, that is adjusted at each iteration using, for example, a BFGS or Symmetric Rank 1 (SR1) update.[5] In a string method, with the gradient perpendicular to the path at each bead i, $\mathbf{g}_i^{\perp}$, given by eq 1, a step direction for bead i is chosen according to

$$\Delta\mathbf{x}_i = -(\mathbf{H}_i^{\perp})^{-1}\mathbf{g}_i^{\perp} \tag{3}$$

where $\mathbf{H}_i^{\perp}$ is a Hessian updated using the gradient perpendicular to the path. Even though eq 3 generates a direction based on a gradient that is perpendicular to the path, there is no guarantee that the direction of $\Delta\mathbf{x}_i$ will also be perpendicular to the path. However, this may not be disadvantageous if (a) the strings' respace operations are able to reverse any path-tangential component of the beads' movements or (b) the tangents of the path do not change significantly during the optimization. Yet, it has been found in the current work that for certain chemical systems, neither (a) nor (b) are satisfied, and that even with respacing, the splines used to represent the path become highly curved, destabilizing the optimization. Another theoretical problem (c) exists with the use of projected forces to construct an approximate Hessian. The resulting Hessian contains curvature information only relating to the directions perpendicular to those through which the tangent has passed; therefore, it becomes inaccurate if the path direction changes significantly.

A number of solutions to the above-described problems may be discussed at this point. An additional projection could be performed to force the step to be perpendicular to the path, i.e.

$$\Delta\mathbf{x}_i^{\perp} = (I - \tau_i\tau_i^T)\Delta\mathbf{x}_i \tag{4}$$

$$= -(I - \tau_i\tau_i^T)[(\mathbf{H}_i^{\perp})^{-1}\mathbf{g}_i^{\perp}] \tag{5}$$

However, this does not solve problem (c) and raises another question. If the tangential component of the step is being projected out, is it necessary to project it out of the gradient as well? Taking a step according to

$$\Delta\mathbf{x}_i^{\perp} = -(I - \tau_i\tau_i^T)[\mathbf{H}_i^{-1}\mathbf{g}(\mathbf{x}_i)]$$

where $\mathbf{H}_i$ is a Hessian updated from nonprojected gradients, solves problem (c), but a new problem emerges, as there is no guarantee that the component of the quasi-Newton step direction, $\mathbf{H}_i^{-1}\mathbf{g}(\mathbf{x}_i)$, that is perpendicular to $\boldsymbol{\tau}_i$ is nonzero. In other words, a step of length zero may be generated by this approach, even if the perpendicular force is significantly nonzero.

All of the above-described issues have been encountered in practice, occasionally on systems as simple as the Müller–Brown PES,[37] which has been a long-serving test case for chain-of-state optimization methods.[10,13,14,29,33] Note that the problem of a quasi-Newton step having a large tangential component does not arise for NEB because the spring forces impart

dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777–786

curvature information into the Hessian that prevents movement of beads toward one another.

3.2. New Optimizer for String Methods. The newly developed optimizer uses a local quadratic approximation to the PES for each bead, as does the QS method.[14] At each iteration, the optimizer moves each bead i, with coordinates $\mathbf{x}_i$, perpendicularly to the tangent τ_i at each bead, subject to the local curvature and trust length. After all beads have been moved, the predesignated spacing criteria are enforced by moving beads along a spline fitted to the new points.

For each bead, i, the local, quadratic approximation to the PES is

$$E_i(\mathbf{x}_i^0 + \Delta\mathbf{x}_i) \approx E(\mathbf{x}_i^0) + \mathbf{g}(\mathbf{x}_i^0)\Delta\mathbf{x}_i + \frac{1}{2}\Delta\mathbf{x}_i\mathbf{H}_i\Delta\mathbf{x}_i \quad (6)$$

Accordingly, each step $\Delta\mathbf{x}_i$ is antiparallel to the gradient perpendicular to the path

$$\Delta\mathbf{x}_i \propto \lambda_i\mathbf{f} := -\lambda_i\mathbf{g}_i^{\perp} \quad (7)$$

The parameter $\lambda_i > 0$ is calculated such that it minimizes the following quadratic expression for the energy, derived by substituting eq 7 into eq 6:

$$E_i(\lambda) = a\lambda_i^2 + b\lambda_i + c \quad (8)$$

where $a = 1/2\mathbf{f}\mathbf{H}\mathbf{f}$, $b = \mathbf{g}(\mathbf{x})\mathbf{f}$, and $c = E(\mathbf{x})$. If the curvature in the direction of $\mathbf{f}$ is negative, i.e., $a < 0$ in eq 8, then there is no minimum and a preset value λ_{max} is chosen.

Having chosen a suitable λ_i, the step is scaled down by a factor $\sigma \in [0,1]$ which depends on the ratio of the actual energy change to the predicted energy change for the previous step. Specifically, a parameter ρ_i is calculated for each bead i as

$$\rho_i = \frac{E_i^n - E_i^{n-1}}{\mathbf{g}_i^n\Delta\mathbf{x}_i^n + \frac{1}{2}\Delta\mathbf{x}_i^n\mathbf{H}_i^n\Delta\mathbf{x}_i^n} \quad (9)$$

where E_i^n, $\mathbf{g}_i^n$, $\mathbf{H}_i^n$, and $\Delta\mathbf{x}_i^n$ are, respectively, the energy, gradient, Hessian, and step for bead i at iteration n, and $\sigma_i(\rho_i)$ is given by

$$\sigma_i(\rho_i) = \begin{cases} \sigma_{min} & \text{if } \rho_i < 0 \\ \min(\sigma_{max}, \sigma_{accuracy}/|1-\rho_i|) & \text{if } \rho_i \geq 0 \end{cases} \quad (10)$$

where $\sigma_{\{min,max,accuracy\}}$ are parameters (given later). Therefore, if $\rho_i < 0$, and the curvature in the direction of the step is the opposite of what is expected, a small step-scale factor is chosen. For a very accurate quadratic energy model, $\rho_i \to 1$ and $\sigma_{accuracy}/|1-\rho_i| \to \infty$, i.e., the maximum step-scale factor, σ_{max}, is chosen. The rational behind such an approach is that if $\sigma_{accuracy}$ reflects some desired maximum deviation of ρ_i from unity, then accuracies in the range $1 - \sigma_{accuracy} < \rho_i < 1 + \sigma_{accuracy}$ deliver the maximum step-scale factor σ_{max}. For accuracies poorer than this, $\sigma_i(\rho_i)$ enters into the regime where the scale factor is significantly reduced.

Hence, the scaled step of bead i is

$$\Delta\mathbf{x}_i = -\sigma_i\lambda_i\mathbf{g}_i^{\perp} \quad (11)$$

If it has any component greater than a parameter Δx_{max}, then it is further scaled so that the maximum component is Δx_{max}. The overall step of the entire string, so far without obeying any spacing criteria, is

$$\Delta\mathbf{X}_{raw} = (\Delta\mathbf{x}_1, ..., \Delta\mathbf{x}_N) \quad (12)$$

As stated above, such a step may perturb the desired spacing, and several approaches to maintain the spacing at each step were tried. If one respaces after every step, it can be difficult to achieve tight convergence, because after every PES-based bead movement, there is a non-PES based movement, which might move the bead in a manner detrimental to the optimization objectives (e.g., uphill). Another possibility is to respace the beads but to avoid potentially detrimental non-PES-based bead movement and perform again a line search along the direction $\mathbf{D} = \mathbf{X}_{respaced} - \mathbf{X}_k$. This approach, however, will generate steps of zero length, causing an optimization to halt, when the bead-wise components of $\mathbf{D}$ are parallel to the beads' tangents, $\boldsymbol{\tau}_i$.

To avoid these issues, beads were respaced only when the fractional spacing error, $\epsilon(\mathbf{X})$, exceeds a user-defined threshold, ϵ_{max}; see the Supporting Information. Thus, the new string vector X_{k+1} is updated as

$$\mathbf{X}_{k+1} = \begin{cases} \mathbf{X}_k + \Delta\mathbf{X}_{raw} & \epsilon(\mathbf{X}_k + \Delta\mathbf{X}_{raw}) < \epsilon_{max} \\ \mathbf{X}_{respace} = \text{Respace}(\mathbf{X}_k + \Delta\mathbf{X}_{raw}) & \text{otherwise} \end{cases} \quad (13)$$

In summary, the new optimization process comprises the following steps:

1. Calculate the gradients $\mathbf{g}(\mathbf{x}_i)$ and energies $E(\mathbf{x}_i)$ (corresponding to $\mathbf{X}_k = [\mathbf{x}_1,...,\mathbf{x}_N]$).
2. If the system has converged, stop.
3. Update the trust parameter (see eqs 9 and 10).
4. Calculate the step $\Delta\mathbf{X}_{raw}$ (see eq 12).
5. Calculate the respaced state vector $\mathbf{X}_{k+1}$ (see eq 13).
6. Repeat from 1.

The QS method[14] also uses a second-order Taylor approximation to the local PES in the vicinity of each bead but then employs a Runge–Kutta integrator to move beads downhill, respacing the beads after every step on the approximate PES, until a trust radius or energy minimum along the search direction is reached. Thus, each global optimization step occurring between electronic structure calculations is actually comprised of many Runge–Kutta steps. Special measures ensure[14] that no bead outruns the others, which could lead to kinks in the splines, especially for very high bead densities. The simplification of the QS method described above is justified since the present work does not aim at obtaining the entire reaction pathway with a very high bead density or very tight convergence criteria.

Determination of convergence is discussed in the next section.

4. COMPUTATIONAL METHODS AND PARAMETERS

All quantum mechanical calculations (evaluation of energies and gradients) were carried out using the program Gaussian 03.[38] The optimization methods and all software were implemented in Python, making use of the numerical/scientific libraries NumPy/SciPy[39] and the Atomic Simulation Environment (ASE).[36]

All local TS searches were carried out using the "Berny" method[6,7,38] of Gaussian 03 with parameters chosen (see the Supporting Information) such that a TS is sought with convergence criteria of $f_{max} = 0.0025$ and $f_{rms} = 0.001667$ hartree Bohr^{-1} or radian^{-1} and additional criteria relating to the predicted step size.[7] No constraints on the number of negative eigenvalues of the Hessian were applied during the optimization. The initial Hessian was calculated analytically and the Hessian of all found TSs was once again calculated analytically.[7] All optimizations were performed at the HF/3-21G level of theory.

dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777–786

This is a lower level of theory but suffices for the development of the current optimization methods.

In the information that follows, distances and Cartesian coordinates are given in Ångströms and angles in radians. Parameters relating to energy (including forces and curvatures) are given in terms of electronvolts (eV).

The convergence of chain-of-state optimizations was judged by examining (1) the root-mean-square (RMS) perpendicular force

$$\mathbf{g}^{\perp}_{\mathrm{RMS}} = \sqrt{\frac{\sum_{i=1}^{N} \left|\mathbf{g}_i^{\perp}\right|^2}{N}}$$

$$N = \text{number of beads}$$

where $\mathbf{g}_i^{\perp}$ is the gradient acting on bead i and (2) the maximum total displacement of the three beads of highest energies over the last three iterations, denoted as x_{total}. In Growing and Searching String methods, a tolerance of $g^{\perp}_{\mathrm{RMS}} < 0.5$ is applied during growth. After the maximum number of beads has been added, a NEB/String is deemed to be converged if

$$g^{\perp}_{\mathrm{RMS}} < 0.1$$

or

$$g^{\perp}_{\mathrm{RMS}} < 0.5 \text{ and } x_{\mathrm{total}} < 0.03 \tag{14}$$

The above convergence criteria for strings or bands containing the final number of beads attempt to minimize the inefficient usage of computer time in the event that the forces are relatively high but the optimizer has reached the limit of its ability to lower them. This is particularly important for NEB optimizations, which, as previously noted,[20,21,33] can be unstable with quasi-Newton-based optimizers, possibly due to there being no simple Lagrangian and nonconservative forces. The optimizer used in the present work takes smaller steps when instabilities occur, i.e., when it is no longer possible to lower the energy. Therefore, the above displacement-based convergence criteria cause the optimization to halt.

For both NEB and string methods, the magnitude of the initial Hessians used by the optimizer was 70*I*. the maximum step size Δx_{max} was 0.1. The parameters σ_{min}, σ_{max} ,and $\sigma_{\mathrm{accuracy}}$ were chosen to be 0.1, 0.9, and 0.1, respectively. The maximum fractional spacing error ϵ_{max} was 0.1.

5. LIBRARY OF TEST REACTIONS

To examine the newly developed methods, the test reactions shown in Scheme 1 were selected. Further details are available in Table 2.

Reaction I, between OH^- and ethane, represents an S_N2 substitution. Reaction II, a rotational isomerism reaction of alanine dipeptide, has been the subject of many previous chain-of-state method investigations.[13−19,23,29] Reactions III and IV are both Diels−Alder reactions between ethene and 1,3-butadiene, the difference being that in reaction III, ethene approaches in the molecular plane of 1,3-butadiene, whereas in reaction IV ethene approaches from outside that plane. These two reactions have previously been used as test cases for related research.[25,29] Reaction V was a hypothesized first propagation step in the transition metal catalyzed polymerization of iso-butene.[40,41] Reaction VI is the addition of water across the double bond of isobutene to form tertiary-butyl alcohol. Reaction VII is another metal-centered propagation step taking place on a highly simplified Ziegler−Natta catalyst.[42]

With two exceptions, all test reactions were studied using both Cartesian and hand-designed Z-matrices coordinates. The exceptions are system V, which was studied using only Cartesian coordinates, and system VII, which was examined using (a) Cartesian and (b) a mixture of Cartesian and Z-matrices coordinates (see section 5 of the Supporting Information, which contains a description of how optimizations were performed using mixed coordinates). For both (a) and (b), the titanium and chlorine atoms were kept fixed; only the two organic fragments were optimized using Cartesian coordinates for (a) and using Z-matrices for (b).

The seven test reactions, examined in various coordinate systems, constitute a total of 13 test cases (Table 2).

Generation of an initial pathway via interpolation between reactants and products using internal coordinates leads to a more physical pathway at lower energy than would be generated from a Cartesian coordinate based interpolation; therefore, for systems where Z-matrices were available, the initial reaction pathway was generated via interpolation using the Z-matrices.[43]

The Supporting Information contains the coordinates of all reactants, products, and TSs and additional discussions on how the reactants and products were optimized.

6. RESULTS

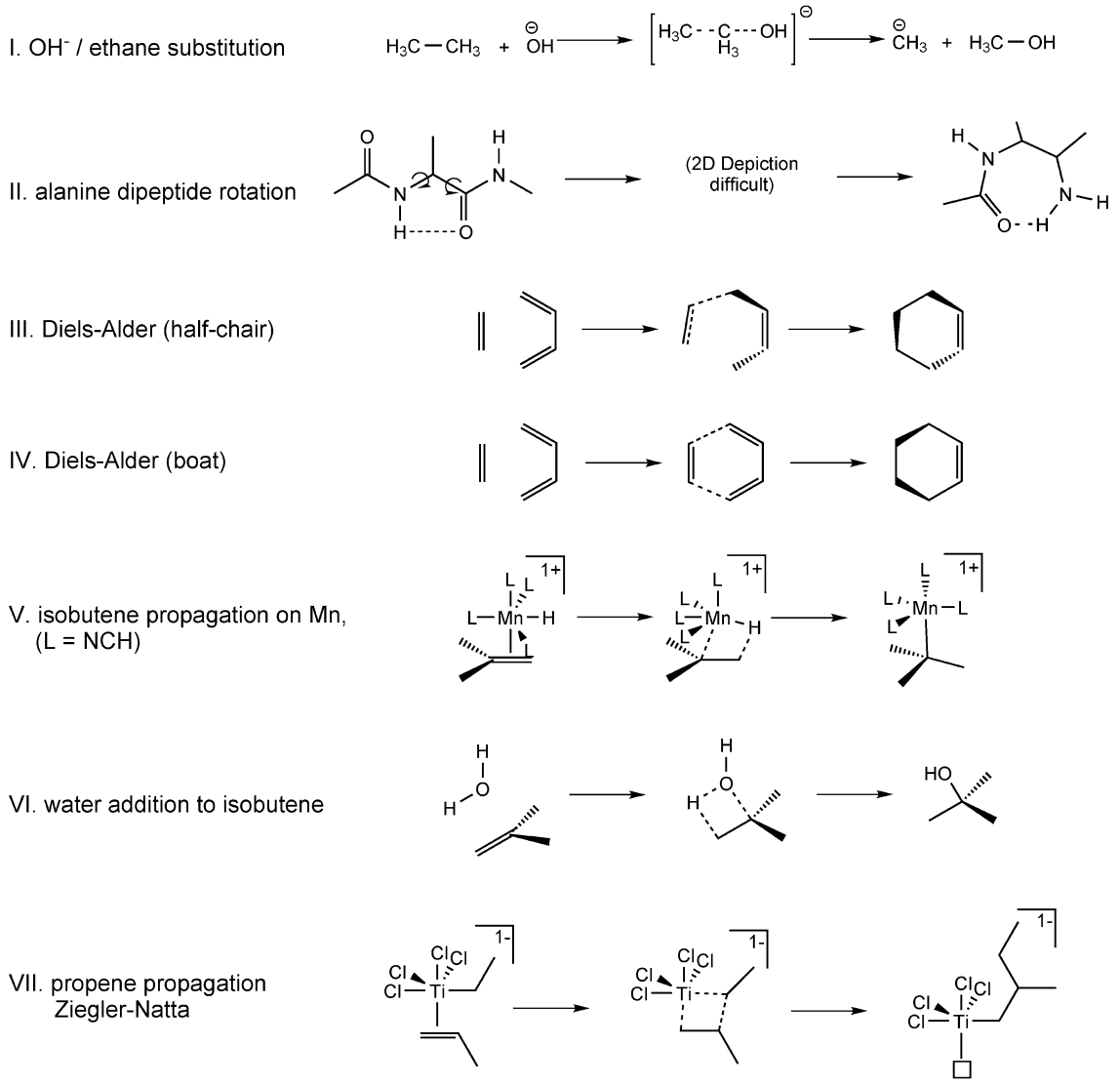
As previously stated, the purpose of the current research has been the topic of TS discovery, rather than determination of the entire reaction pathway. Consequently, we are interested in chain-of-state optimizations with a number of beads that lies close to the transition between success and failure of the subsequent local transition state search. The number of beads that this corresponds to was found to be close to 6.

6.1. String/Band Convergence. The number of gradient calls required for convergence of the herein reported Searching String method has been compared to that of the NEB, String and Growing String methods. For 6-bead searches of each test case, this information is presented in Figures 3 and 4. Figure 3 contains the raw number of gradients required by each method for each test case, and Figure 4 contains the aggregated, normalized gradient counts for all test cases. The normalized gradient count was calculated as the gradient count divided by the average gradient count for a particular system, allowing both difficult and easy problems to be given an equal weighting when considering the overall performance of a method.

From Figure 3 it can be seen that certain test cases have an inherently greater complexity. For example, irrespective of the coordinate system or method that is used, test reaction VII always requires significantly more gradient calls for convergence than test reaction II. Also, each of the methods have a test case on which they perform relatively poorly. For this reason, the aggregated, normalized gradient call results in Figure 4 have been prepared. The results depicted therein (and also those in Figure 3) suggest that, on average, the Searching String method is able to converge a pathway with a far higher bead density in the region of interest but with a number of gradient calls similar to other string methods with the same number of beads. The bead density for a six-bead chain-of-state method with even spacing is 5, whereas application of the Searching String method with the same number of beads leads to a maximum density of 12 (see Table 3).

6.2. Interpolation Methods. Having examined the convergence behavior of the Searching String method, the effectiveness of the five interpolation methods summarized in

Scheme 1. Reactions upon Which All Methods Have Been Tested (See Also Table 2)[a]



[a]The Supporting Information contains the coordinates of all structures and details of the calculation.

Table 2. Additional Information Pertaining to the Test Reactions Shown in Scheme 1[a]

ID	reaction	charge	spin multiplicity	coordinates[b]
I	OH^- addition to ethane	−1	1	cart, zmt
II	alanine dipeptide	0	1	cart, zmt
III	Diels–Alder (half chair)	0	1	cart, zmt
IV	Diels–Alder (boat)	0	1	cart, zmt
V	isobutene propagation on Mn	1	6	cart
VI	water addition to isobutene	0	1	cart, zmt
VII	propene propagation, Ziegler–Natta	−1	1	cart, mixed

[a]Due to the use of Cartesian (cart), Z-matrix (zmt), and mixed coordinates for most systems, a total of 13 test cases were derived from the seven reactions. [b]Types of coordinates used for optimization; see Supporting Information.

Table 1, that aim to generate an improved TS estimate, is now explored.

It would be ideal to be able to judge the quality of a TS estimate without resorting to a full local TS search. One possibility for such an *a priori* test might in principle be based upon the forces acting on the atoms of an estimated TS (i.e., prior to local optimization). This measure, however, has been found to be a poor predictor of a TS estimate's quality: very often, estimates with high initial forces would very favorably converge to the correct TS, with the opposite also being observed. The measure of quality of the TS estimate that has the most practical relevance is the number of gradient calls taken by a local optimization algorithm to converge to the correct TS.

As an indicator of success, it is not possible to use the number of gradient calls in a local TS search directly, or after averaging them, since some estimates will lead to a TS search that either never converges or converges to the wrong TS, despite the fact that only limited compute time might have been required to achieve the false result. Furthermore, not all test cases led to local searches with the same inherent expense—some test cases require local searches typically of less than 10 gradient calls, while others require local searches typically in the range of 30–40 gradient calls. Therefore, it is desirable to devise a scoring mechanism that does not favor any particular interpolation method if it happens to perform well for one of the

dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777–786

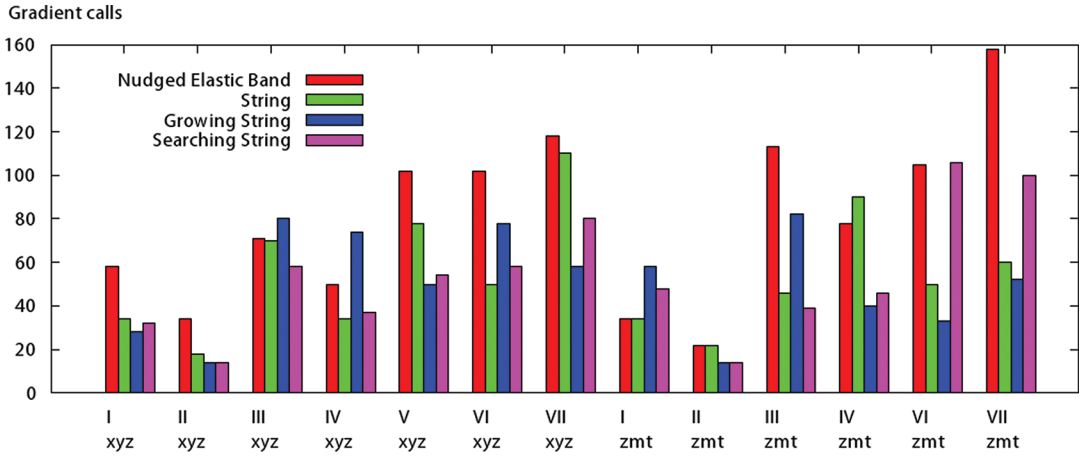


Figure 3. Benchmarking of the Nudged Elastic Band, String, Growing String, and Searching String methods with six beads, applied to 13 test cases which were generated from the seven exemplary reactions introduced in Table 2. The optimization was done in Cartesian coordinates (xyz) or Z-matrix coordinates (zmt). Bar height represents the number of gradient calls required to converge the string or band (lower is better).

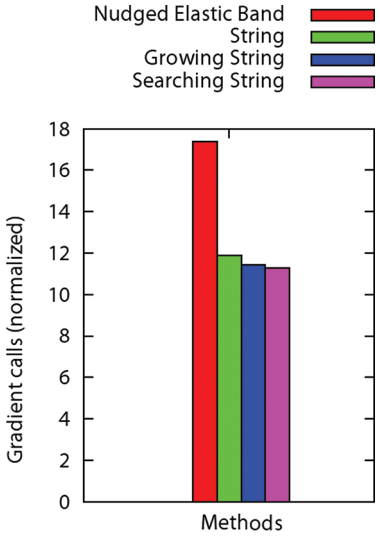


Figure 4. Benchmarking of the Nudged Elastic Band, String, Growing String, and Searching String methods with six beads, applied to 13 test cases derived from seven test reactions with various coordinate systems (Table 2). Bar height represents the (normalized) total count of gradient calls (see text) required to converge the string or band for the 13 test cases (lower is better).

Table 3. Chain-of-State Methods and the Number of Beads Used for the Selected 13 Optimization Problems (Table 2)[a]

chain-of-state method	beads	max. bead density[b]
Nudged Elastic Band	4	3
(Growing/Searching) String	4	3
Nudged Elastic Band	5	4
Growing String	5	4
Searching String	5	6
Nudged Elastic Band	6	5
Growing String	6	5
Searching String	6	12

[a]Note that Growing, Searching and normal String methods with 4 beads are trivially equivalent, as no growth is possible. [b]The maximum bead density is defined as the reciprocal of the minimum interbead distance as a fraction of the total string length.

expensive systems. Thus, the following scoring mechanism was chosen (higher is better):

- A separate score between 0 and 1 is awarded to each application of each interpolation method to each test case converged by each chain-of-state method.
- With the minimal and maximal numbers of gradient calls, $grads_{min}$ and $grads_{max}$, we define the score as $grads_{min}$/grads for each test case of a given method that ultimately converges. Thus, the values of the score vary between $grads_{min}/grads_{max}$ and 1, with 1 being awarded by definition to the case that converges fastest.
- An interpolation method which generates a TS estimate for which the local search does not converge, or one which converges to something other than the desired transition state, is given a score of zero.

The 13 test cases were examined using an assortment of eight NEBs, Strings, Growing Strings, and Searching Strings (see Table 3). From each of the 8 × 13 = 104 converged NEBs and strings, five TS estimates were generated and used as input for a local TS search. The success of each local search was converted into a score as described above, and the cumulative scores for each interpolation method for the eight chain-of-state optimizations appear in Figure 5. Each vertical bar in Figure 5 thus has a height equal to the sum of 13 scores from each of the 13 test cases. Each group in that figure contains results for a particular NEB or String, and the columns within each group correspond to a particular interpolation method.

The application of the chain-of-state method that resulted in the poorest local searches, that is, the one that had the poorest ability to locate the TS, was the NEB method with only four beads. The method which exhibited the best behavior, as judged only by the successes of the local TS search, was the Searching String method. Importantly, the worst interpolation method in every case was the method of choosing the highest bead. Thus, any interpolation method can be used confidently, without fear of obtaining a poorer TS estimate than would be delivered by the highest bead.

A slightly surprising result is that for the NEB and Growing String methods, when going from five to six beads, the achieved score decreases. By examining the convergence behavior of individual test reactions, this has been found to be for two reasons. First, many systems appear to benefit from having a

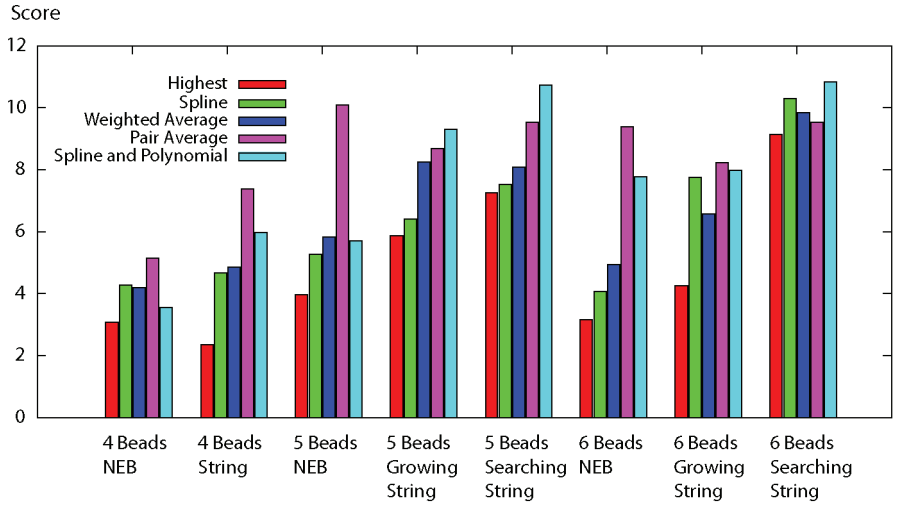


Figure 5. Benchmarking of interpolation methods. The height of each vertical bar represents the sum of 13 scores (higher is better). Each group contains results for a particular chain-of-state method and the columns within each group correspond to a particular interpolation method.

bead exactly in the middle, because this is close to where the TS typically occurs along the reaction path. Second, in the Growing String method with five and six beads, a slightly more subtle effect is operative. In applications of the Growing String method with five or six beads, a single growth step takes the strings from four to five and four to six beads, respectively. For the string of six beads, the geometries of the two new beads are generated from existing beads that are further from the TS than in the string of five beads. This has led to middle beads of higher energy, even though they were converged to the same level. However, with arbitrarily tight convergence criteria, Growing Strings of both five and six beads should converge to a similar reaction energy profile. While in general it is expected that a large-scale increase in the number of beads will lead to a more accurate TS estimate, this is not guaranteed for smaller changes, with similar observations made in other research on the growing string method.[13] The fact that this inverse relationship between performance and the number of beads has not been observed for the Searching String method highlights a particular strength: if an optimization with N beads fails to yield a usable TS estimate, then upon continuing to $N + 1$ beads, (a) beads will not be inadvertently moved away from the TS and (b) the computational effort involved in converging the N bead Searching String can be salvaged as the starting point for the search with $N + 1$ beads.

For the four-, five-, and six-bead NEB applications, the interpolation method that stands out as superior is the *Pair Average*, for which the geometry of the estimated TS is calculated as a linear average of the beads whose path-parallel energy gradients indicate that they bracket the TS. The NEBs and the four-bead String method are the only ones where the *Pair Average* method stands out as significantly better than the others, and this may be rationalized as follows. On the one hand, the NEB method is the only one which does not use spline-based tangents; therefore, interpolations using splines may be less relevant. On the other hand, the four-bead String, while utilizing spline-based tangents, has so few beads that the beads on either side of the TS are so far from the TS that PES and tangent information at these points is less relevant than for beads very close to the TS.

The method with the greatest bead density, the Searching String with six beads, enjoyed the smallest gains from application of TS interpolation methods as compared with choosing the highest bead (Figure 5). This is expected since a greater bead density in the vicinity of the TS means that the highest bead is more likely to be close to the correct TS. The method with the second greatest bead density, namely the Searching String with five beads, also exhibits the second best performance and represents a significant improvement even upon the best Growing String.

7. CONCLUSION

Improvements in TS discovery using discretized reaction pathways (also known as chain-of-state methods[1]) have been developed. Specifically, this work focused on two aspects: (1) development of the "Searching String" method, which searches actively for the TS containing region of the pathway during convergence, and (2) methods of generating more accurate TS estimates from an optimized String or Nudged Elastic Band (NEB). The developed methods have been benchmarked on 13 test cases covering a wide variety of chemistries (e.g., unimolecular, bimolecular reactions, reactions of metal complexes) and coordinate systems (Cartesian and internal coordinates and mixtures thereof).

Development 1 involved the utilization of gradient and tangent information to determine which pair of beads on a discretized reaction pathway was likely to bracket the best TS estimate. On the basis of this information, a bead is added to the midpoint of the TS-containing region, with the process being continued until some maximum number of beads is reached. This approach has the favorable outcome that, due to the higher bead density in the TS containing region, such optimizations more frequently lead to TS estimates which converge efficiently during a subsequent local TS optimization. Importantly, this advantage does not come at the expense of the efficiency with which the string converges. An additional strength of the Searching String method is that, contrary to the situation with an even bead spacing, an N-bead string is a starting point for a $(N + 1)$-bead one, where (a) the computer time invested in the convergence of the former is beneficial for convergence of the latter and (b) it is unlikely that an increase in the bead density moves beads away from the best TS guess.

Development 2 involved investigation of five methods of extracting a TS estimate from a converged chain-of-state. The simplest of these methods involves the use of the highest

dx.doi.org/10.1021/ct200639w | J. Chem. TheoryComput. 2012, 8, 777−786

energy bead as the TS estimate and the most sophisticated uses a cubic polynomial fitted to the energy profile. All investigated interpolation methods generated a more favorable TS estimate as compared to choosing the bead with the highest energy, with the cubic-polynomial-based method featuring the greatest improvement.

■ ASSOCIATED CONTENT

Supporting Information

Details of the implementation; raw data used to produce Figures 3 and 5; geometries of all reactants, transition states, and products. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: roesch@mytum.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

H.C.-M. acknowledges a postdoctoral grant from the Alexander von Humboldt Foundation. A.N. is grateful for support by the International Graduate School of Science and Engineering at Technische Universität München. S.K. acknowledges support by Bundesministerium für Wirtschaft und Technologie, project no. 02E10186. We also acknowledge generous computing resources at Leibniz Rechenzentrum München. This work was supported by Fonds der Chemischen Industrie.

■ REFERENCES

(1) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; Wiley: Chichester, England, 2006.
(2) Weinan, E.; Vanden-Eijnden, E. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391−420.
(3) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **1999**, *111*, 7010−7022.
(4) Klimes, J.; Bowler, D. R.; Michaelides, A. *J. Phys.: Condens. Matter* **2010**, *22*, 074203.
(5) Nocedal, J.; Wright, S. *Numerical Optimization*, 2nd ed.; Springer: New York, 2006, Springer Series in Operations Research and Financial Engineering.
(6) Schlegel, H. B. *J. Comput. Chem.* **1982**, *3*, 214−218.
(7) Gaussian Website. http://www.gaussian.com. (accessed August 19, 2011).
(8) Jónsson, H.; Mills, G.; Jacobsen, K. W. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: River Edge, NJ, 1998; Chapter Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions.
(9) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(10) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978−9985.
(11) Weinan, E.; Ren, W.; Vanden-Eijnden, E. *Phys. Rev. B* **2002**, *66*, 052301.
(12) Weinan, E.; Ren, W.; Vanden-Eijnden, E. *J. Chem. Phys.* **2007**, *126*, 164103.
(13) Peters, B.; Heyden, A.; Bell, A. T.; Chakraborty, A. *J. Chem. Phys.* **2004**, *120*, 7877−7876.
(14) Burger, S. K.; Yang, W. *J. Chem. Phys.* **2006**, *124*, 054109.
(15) Burger, S. K.; Liu, Y.; Sarkar, U.; Ayers, P. W. *J. Chem. Phys.* **2009**, *130*, 024103.
(16) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *130*, 244108.
(17) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *Chem. Phys. Lett.* **2010**, *484*, 392−398.
(18) Quapp, W. *J. Chem. Phys.* **2005**, *122*, 174106.
(19) Quapp, W.; Kraka, E.; Cremer, D. *J. Phys. Chem. A* **2007**, *111*, 11287−11293.
(20) Maragakis, P.; Andreev, S. A.; Brumer, Y.; Reichman, D. R.; Kaxiras, E. *J. Chem. Phys.* **2002**, *117*, 4651−4658.
(21) Chu, J.-W.; Trout, B. L.; Brooks, B. R. *J. Chem. Phys.* **2003**, *119*, 12708−12717.
(22) Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *120*, 2082−2094.
(23) Koslover, E. F.; Wales, D. J. *J. Chem. Phys.* **2007**, *127*, 134102.
(24) del Campo, J. M.; Köster, A. M. *J. Chem. Phys.* **2008**, *129*, 024107.
(25) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1997**, *107*, 375−384.
(26) Dey, B. K.; Janicki, M. R.; Ayers, P. W. *J. Chem. Phys.* **2004**, *121*, 6667−6679.
(27) Burger, S. K.; Ayers, P. W. *J. Chem. Phys.* **2010**, *132*, 234110.
(28) Burger, S. K.; Ayers, P. W. *J. Chem. Theory Comput.* **2010**, *6*, 1490−1497.
(29) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *129*, 174109.
(30) Bettens, R. P. A; Collins, M. A. *J. Chem. Phys.* **1999**, *111*, 816−826.
(31) Dawes, R.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. *J. Chem. Phys.* **2008**, *128*, 084107.
(32) Jin, C. M. *Commun. Comput. Phys.* **2007**, *2*, 1220−1243.
(33) Sheppard, D.; Terrell, R.; Henkelman, G. *J. Chem. Phys.* **2008**, *128*, 134106.
(34) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. *Phys. Rev. Lett.* **2006**, *97*, 170201.
(35) Liu, D. C.; Nocedal, J. *Math. Program.* **1989**, *45*, 503−528.
(36) Bahn, S. R.; Jacobsen, K. W. *Comput. Sci. Eng.* **2002**, *4*, 56−66.
(37) Müller, K.; Brown, L. D. *Theor. Chim. Acta (Berl.)* **1979**, *53*, 79−93.
(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc., Wallingford CT, 2004.
(39) SciPy: Open Source Scientific Tools for Python. http://www.scipy.org (accessed November 8, 2011).
(40) Chaffey-Millar, H.; Kühn, F. E. *Appl. Catal., A* **2010**, *384*, 154−164.
(41) Nuyken, O.; Vierle, M.; Kühn, F. E.; Zhang, Y. M. *Macromol. Symp.* **2006**, *236*, 69−77.
(42) Cossee, P. *J. Catal.* **1964**, *3*, 80−88.
(43) Thompson, H. B. *J. Chem. Phys.* **1967**, *47*, 3407−3410.